The London School of Economics

and Political Science

EMPIRICAL ESSAYS ON THE INTERACTION BETWEEN HOUSING AND LABOUR MARKETS

Sejeong Ha

A thesis submitted to the Department of Geography and Environment of the London School of Economics for the degree of Doctor of Economic Geography, London, December 2012

DECLARATION

I certify that the thesis I have presented for examination for the MPhil/PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of 43,417 words.

I can confirm that Chapter 4 is partially based on my MSc Dissertation at London School of Economics. I can confirm that my thesis was copy-edited for conventions of language, spelling and grammar by Darren Southcott.

ABSTRACT

This thesis presents three empirical essays on the interaction of housing and labour markets, which generates academically meaningful social outcomes. The first essay looks at whether one's tenure choice affects unemployment as this question has potential implications for homeownership subsidy schemes adopted in many advanced countries. The contribution of this essay is mainly methodological in that it rigorously deals with the endogeneity of homeownership by taking an IV approach with instruments not adopted previously for studies in the UK in conjunction with panel data models. Using the local homeownership rate and parental homeownership status as an instrument, it shows that homeownership does not increase the probability of being unemployed.

The second essay highlights the role of local housing market information as a determinant of housing tenure. As the distance a mover wants to move increases, the costs of collecting information on the destination housing market rise and the quality and amount of the information collected fall. Therefore, it is hypothesised that the longer the distance moved, the more likely movers are to choose private renting over owner-occupation since homeownership decisions require a large amount of information on the target properties and their neighbourhoods. Empirical tests that control for relevant characteristics correlated with distance moved and tenure decisions provide supporting evidence for this hypothesis.

The last essay is the first UK study to confirm that commuting time has a negative influence on worker effort. The topic has important implications for transportation

policy, employer's commuting welfare strategy and hiring decisions and individual worker's location decisions. As commuting is physically and mentally tiring, it could influence worker effort negatively. The hypothesis turns out to be true when the absenteeism rate and unpaid overtime hours are used as proxy variables for work effort.

ACKNOWLEDGEMENT

First and foremost, I would like to express my sincere gratitude to my supervisor, Prof. Paul Cheshire and review supervisor, Dr. Christian Hilber for accompanying, encouraging and supporting me patiently over the long journey I could not have finished without them. Their critical but inspiring advice and comments have helped me to thread my premature ideas into this thesis. Needless to say, the invaluable comments I received at seminars and conferences were of great help in forming this thesis, too.

I have been so lucky over the course of the PhD to meet Koun, Sujung, Youngjeong, Hyunbang, Jaejin, June, Filipe, Gerard, Fran and Frank with whom I spent the most memorable times at LSE. My heartfelt thanks also goes out to London Moim members, my spiritual family, for comforting and inspiring me during the most challenging time of my life. I wish I could express my special thank to Mr. Leonard Biles in person but sadly he is no longer around.

I would like to thank my parents, Mr. Il-nam Ha and Ms. Soon Rye Kang for unlimited support, love and patience. I am also thankful to my brother, Seyoon and his wife, Hongju for caring for my parents alone in my absence for such a long time. Finally, I am greatly indebted to my wife, Sun Mi for raising our son, Jeongwoo without much help from me, when she needed it most.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

1.1. Overview of thesis

This thesis explores the interaction between housing and labour markets from an economic perspective through three empirical essays. The two markets are distinguishable from each other in terms of elements and functioning and have been studied separately by two distinctive groups of economists with different academic interests and focuses. As a result, two branches of applied economics, namely urban economics (more specifically housing economics) and labour economics, have developed. Housing economics particularly tends to focus on institutional arrangements such as the provision of social housing or the specifics and role of housing finance whereas labour economists tend to look at labour markets in the aggregate and certainly in a non-spatial way.

However, the spatial extent of both housing and labour markets is similar for most households as both work and other daily activities need to occur within a relatively small area for the efficiency of time and resource. The spatial proximity of the two markets increases the likelihood of socially and economically meaningful interaction between them. For example, the location of the job determines that of the residence and, at the same time, the location of residence tends to determine the spatial area of the search for jobs. The starting point of this thesis is the recognition of the fact that spatial dimension is exactly what housing and labour markets have in common and therefore it is difficult to understand housing markets without taking into account the spatial character of the labour market and vice versa.

There are a few academic advantages of studying the two markets jointly. The studies of the interaction between the two markets are usually initiated by one side and they soon attract the attention of the other. Through the process of competition and collaboration between the two sides, the expert knowledge and research specialities of one side are likely to be mixed with those of the other. For instance, the research focus of housing economists could inspire labour market researchers to pay more attention to spatial aspects of the labour market such as the locations of jobs and workers. Any critical errors which might threaten the validity of the studies can be more easily detected and corrected. Indeed, labour economists often test with macro-data Oswald's (1996) hypothesis that one's homeownership may affect one's employment status but urban economists criticise this practice and employ micro-data as the hypothesis is concerned with behaviours and labour market outcomes of *individual homeowners*.

Looking at one of the two markets can also help in finding missing pieces of the puzzle of the other market. Although there are many studies which do not confirm Oswald's (1996) hypothesis, his speculation could have led to the discovery of a significant cause of unemployment from a completely unexpected angle. Moreover, larger scale and more complete pictures of the operation of urban areas can be captured through the study of the two most notable urban markets simultaneously. As they interact to generate joint social outcomes, an observation of only one market is not able to reveal a complete picture of how urban areas work and operate. For example, to study commuting behaviours, the functioning

of labour, housing and transportation markets need to be addressed together as Rouwendal and Nijkamp (2004) suggest.

This thesis tries to explore three topics, all of which are related to unemployment, one of the most important economic issues in our society. Each chapter commonly focuses on activities or behaviour to find and retain jobs and studies the relationships between such activities and socio-economic outcomes in labour or housing markets. Chapter 2 focuses on individual's homeownership decisions to ascertain if it causes unemployment. Given the greater cost of inter-city mobility faced by homeowners compared to private renters, it would be more difficult for homeowners to move labour markets and so, they may experience a greater probability of being unemployed. Chapter 3 looks at the relationship between migration distance and housing tenure decisions. Given the costs of information on more distant housing markets, an inter-city/labour market move would make a worker more likely to opt to rent rather than own in their destination housing markets. Chapter 4 is concerned with commuting which is an alternative mechanism to migration to find and keep a job. Within a city's labour market given the costs of commuting, a longer journey to work is likely associated with less effort on the job. A more detailed discussion of each chapter is given below.

1.2. Outline of empirical chapters

1.2.1. Homeownership and unemployment

Chapter 2 investigates whether one's homeownership causally affects one's employment status. Homeownership is subsidised through various policies in many countries but if it causes unemployment, one of the most important economic problems many governments struggle to resolve, those policies may have to be revised or even abandoned. Therefore, it is important to understand any link between homeownership and unemployment for public policy's sake. One well-known hypothesis on the relationship between homeownership and unemployment is speculated first by Oswald (1996) who shows that homeownership rates are positively correlated with unemployment rates across countries. This suggests that the effect from homeownership to unemployment might be causal since homeowners are generally less mobile than renters given the costs of buying and selling housing. Initially, most studies tested the hypothesis using macro-data and almost unanimously produced empirical results in support of Oswald's proposition (e.g. Nickell, 1998; Pehkonen, 1999; Green and Hendershott, 2001). However, the criticism arises that they used macro-data to explain a micro-level outcome (note that *individual* homeowners' low mobility is suggested as a cause of their unemployment). More recent studies employ micro-data and generally produce evidence which contradicts the conclusions of the macro-level studies (e.g. Van Leuvensteijn and Koning, 2004; Munch et al., 2006, 2008; Battu et al., 2008; Coulson and Fisher, 2009).

Furthermore, those micro-level studies tend to deal with the endogeneity of homeownership much more rigorously than the macro-level studies by often

adopting the multi-spell approach (e.g. Van Leuvensteijn and Koning, 2004; Munch et al., 2006, 2008, Battu et al., 2008). The endogeneity of homeownership is problematic as it causes ordinary least squares (OLS) or maximum likelihood (ML) estimators to be biased for two reasons. Firstly, one's homeownership status is correlated with one's personal and family characteristics and these are simultaneously correlated with one's unemployment. The omission of the variables representing those characteristics from OLS regressions would cause bias. Secondly, there is mutual causality between homeownership and unemployment. Oswald (1996) suggests that homeowners' low mobility could lead to a higher likelihood of being unemployed at any given time but reversely, the unemployed are less likely to be homeowners as employment is an important source of income to finance home purchases. Therefore, the OLS or ML estimation would capture the correlation between homeownership and unemployment rather than the causality this chapter focuses.

This chapter also explores the relationship between homeownership and unemployment using micro-data but it is methodologically distinguished from the previous micro-level studies in that it relies on an instrumental variable (IV) approach rather than the multi-spell approach. This chapter employs two instruments but no single UK dataset can provide both of them at the same time. Therefore, these two instruments are obtained from different datasets and an empirical analysis needs to be carried out for each instrument separately. The first instrument is the local homeownership rate which can be regarded as a comprehensive indicator of how accessible homeownership is in the local area. It is a frequently employed instrument for individual homeownership status (e.g.

Van Leuvensteijn and Koning, 2004; Dujardin and Goffette-Nagot, 2008) but has not yet been tried in the UK study of this topic. One of the special licensed versions of the British Household Panel Survey (BHPS) provides information on the survey respondent's local district of residence and based on this locality information and housing tenure status, homeownership rates of 278 districts of Great Britain are calculated.

The second instrument is parental homeownership status when individuals were young children. Parental homeownership status was once used as an instrument for the Danish study by Munch et al. (2008) and shown to be highly correlated with their offspring's homeownership status. The high correlation is also found in the related literature, which is believed to be because of an inter-generational transfer of wealth and knowledge of real estate transactions (Munro, 1989; Hamnett et al., 1991; Deutsch, 1997; Haurin and Morrow-Jones, 2006; Hilber and Liu, 2008). However, parents' homeownership status is unlikely to be correlated with their children's unemployment status. This instrument is available from the British Cohort Study (BCS) 1970 which follows cohort members born in a certain week of 1970 throughout their lifetimes. Therefore, the information on the homeownership status of their parents can be traced back to when the cohort members were 5 years old.

The model specification states that one's unemployment status is a function of homeownership status along with other controls. The model is estimated first using the ML logit by pooling the entire sample over time from the BHPS. Contrary to Oswald's hypothesis, the likelihood of being unemployed turns out to

be lower for homeowners than for private renters but it is likely to be a spurious correlation through individuals' innate abilities which reduces the likelihood of unemployment but increases that of homeownership. To control for unobserved time-constant heterogeneity including innate ability and to remove the possibility of reverse causality, a fixed effects (FE) IV model is estimated with the instrument being the local homeownership rate. The result indicates no causal effect of homeownership on unemployment status. When the model is estimated by two-stage least squares (2SLS) with parents' homeownership status as an instrument using the BCS 1970, the same result is found – homeownership has no impact on unemployment status.

1.2.2. Distance moved and housing tenure decision

Chapter 3 focuses on the effect on movers' tenure choices of information on destination housing markets, with the distance moved being a proxy variable for the costs of acquiring the market information. It is often observed that job-related movers move relatively long distances as they are likely to move between spatially separated labour markets. Such inter-market moves naturally lead to a change in home location and housing markets, too. In the new housing markets, the movers face uncertainty about the quality of the neighbourhood and the appropriate price level for housing of a given quality. This uncertainty is expected to cause the movers to opt to rent since renting implies lower subsequent moving costs and no house price risk that can be substantial for uninformed movers. Once renters gather more information on their housing markets over time, they can make more informed decisions regarding the location of more permanent housing. Yet, the degree of uncertainty is expected to increase with the distance moved as

the collection of information on the new housing market becomes increasingly difficult as movers intend to move longer distances. Hence, it can be hypothesised that the longer the distance moved, the higher the likelihood of choosing private renting over owner-occupation.

The dataset for this study is the Survey of English Housing (SEH) as it contains the essential information such as the distances moved to current accommodation and housing tenure status of households. In addition, various socio-demographic and economic characteristics of household heads are available to be included as control variables. These variables are of great help for reducing omitted variable biases as they are thought to be correlated with distance moved and propensity to own. Furthermore, though it is a cross-sectional dataset, the SEH provides some information on the pre-move conditions of households including previous housing tenure which helps to control for unobservable preferences and ability to own.

Including the main hypothesis introduced above, three hypotheses, all of which are consistent with the notion that the distance moved affects a tenure decision, are tested. The first hypothesis is whether the amount of information on destination housing markets held by movers decreases with the distance moved. The measures of how aware movers are of problems in the new neighbourhoods are assumed to reflect the amount of local housing market information held by the movers prior to their moves. Empirically, the degree of awareness is shown to decrease with the distance moved. Secondly, the main hypothesis that the likelihood of choosing homeownership over private renting declines with the distance moved is shown to have a support from the empirical analysis. The last

hypothesis is concerned with household's subsequent moving decisions and states that the longer the distance moved, the more likely movers are to move again shortly after their initial moves. As renting households become familiar with their housing markets, they may consider moving to more permanent accommodation within the same market. This tendency is likely to be stronger for longer-distance movers as they are less likely to be satisfied with their new neighbourhood and accommodation due to lack of information on them. Therefore, they are expected to move out sooner than shorter-distance movers and this hypothesis is confirmed by the empirical analysis.

1.2.3. Commuting time and worker effort

Chapter 4 examines whether commuting time has a negative influence on worker effort. Though inevitable in the modern world, commuting is recognised as a physically and mentally tiring activity. Psychology and transportation studies also report commuting as a negative experience (Koslowsky et al., 1995; Evans and Wener, 2006; Hoehner et al., 2012). This finding naturally leads to the question of whether the length of commuting time affects worker effort negatively. Though there could be important implications for a study of this kind to workers, employers, and policy makers, economists have not paid very much attention to the question to date. Zenou (2002) uses the idea of commuting distance reducing worker effort levels only to set up his red-line hypothesis which seeks to explain Kain's (1968) spatial mismatch hypothesis and it is not tested empirically. Ha (2005) tests the relationship using the UK data but cannot find evidence for it. To the best of the author's knowledge, the paper by Ommeren and Gutiérrez-i-Puigarnau (2011) is the first and only study that carries out more rigorous research

on this issue and concludes that commuting distance does increase the worker absenteeism rate using German data. Given this relative lack of information and evidence, this chapter contributes to the literature by adding evidence using UK micro-data.

In the literature, effort is understood as effective labour supply and this allows the use of a neoclassical labour supply model as an analytical framework to predict the effect of commuting time on effort. The basic idea is that a worker facing reduced leisure time as a result of more time spent commuting is likely to shirk work to achieve the same utility level as an otherwise identical worker with zero commuting time. Through the review of the labour-supply literature, absenteeism and unpaid overtime work hours are chosen as proxy measures for effort. Though they are correlated with effort in different ways (absenteeism is negatively and unpaid overtime work is positively correlated), both variables bear a core characteristic of effort: effort is costly to employees but beneficial to employers. The adoption of two different measures is expected to strengthen the reliability of the empirical results.

The dataset for empirical analysis is the UK Quarterly Labour Force Survey (QLFS) as it provides information on both absenteeism and unpaid overtime work hours and other relevant variables which contain a great deal of information relating to work and employment. The availability of abundant information on personal and family characteristics and work conditions helps to control all the necessary explanatory variables so that concern about omitted variable bias can be substantially lessened. The empirical results are supportive of the hypothesis:

absenteeism is positively correlated with commuting time and unpaid overtime work hours are negatively correlated.

To check the robustness of the results, this chapter examines if the size of the effect is different between full-time and part-time workers. The prediction is that the relationship between commuting time and effort level would be stronger for part-time workers because the effective wage rate would fall more sharply for them since their hours of work are shorter for a given increase in commuting time. As predicted, part-time workers reduce their effort levels more substantially for a given increase in commuting time. Secondly, working women are expected to reduce effort at a greater degree than men as commuting time increases. They are likely to be more time-constrained than working men due to additional household work and so, they are expected to respond more sensitively to a given increase in commuting time. This prediction is also shown to hold true empirically. Even after finding evidence consistent with the hypothesis that commuting time affects worker effort negatively, one cannot be certain that the confirmed correlation is entirely causal. Theoretically, reverse causation running from effort to commuting could exist as less work-oriented workers might choose long commutes to live in suburban areas to enjoy a higher quality of life. However, less work-motivated workers are likely to try to find jobs near their homes and therefore the influence of the reverse causality is likely to be limited.

The empirical findings of this chapter have important implications for policy-makers, employers and workers. For policy-makers, this chapter suggests that increased labour productivity should be considered as one of the benefits of

improved transportation infrastructure in the cost and benefit analysis of investment in public transport. Employers would find it worth improving their workers' commuting environment or shortening their commuting time by, for example, running free shuttles or subsidising the use of faster and more convenient transport, since all such measures can increase worker effort levels and productivity. For workers, an increase in their effort levels due to reduced commuting time could lead to pay rises and promotion through increased productivity. They may consider moving closer to their workplaces if the gain from improved productivity is predicted to be greater than the relocation costs.

CHAPTER 2. HOMEOWNERSHIP AND UNEMPLOYMENT

2.1. Introduction

Assessing the social benefits and social costs of homeownership accurately is important from a welfare economics point of view as the social *net* benefit (or cost) determines whether homeownership should be encouraged and subsidised. In reality, many advanced countries provide favourable tax treatment for homeowners (Hendershott and White, 2000) and this may be because the relative benefits of homeownership have been demonstrated and recognised, such as its positive impact on social capital, child-rearing, home maintenance, etc. However, the debate still continues around whether homeownership has an unfavourable impact on one's labour market outcome. This chapter aims to provide rigorous and novel evidence on the causal effect of homeownership on unemployment using UK micro-data and complements earlier studies in that it is the first to combine a panel fixed effects model and an instrumental variable approach to assess the causal effect.

The recent revival of the debate on the impact of homeownership on unemployment is sparked by Oswald (1996, 1999) as he suggests that persistently high unemployment rates have been associated with rising homeownership rates in most advanced economies since the 1960s. Early empirical studies test whether those two rates have a positive correlation at the country and regional level and confirm it almost unanimously (e.g. Nickell, 1998; Pehkonen, 1999; Green and Hendershott, 2001). However, these studies are criticised as they depend on overly simplistic econometric techniques and control for too few explanatory variables to

cope with the endogeneity of homeownership. A more fundamental problem of these studies arises from the fact that they suggest *macro-level* findings as evidence for Oswald's hypothesis whereas it is concerned with *individual* homeowners' behaviours and labour market outcomes. Therefore, it is questionable whether aggregate-level results are truly brought about by underlying individuals' behaviours.

Recent micro-level studies tend to look directly at the unemployment status of individual homeowners and conclude that they are not necessarily more likely to be unemployed than renters (e.g. Van Leuvensteijn and Koning, 2004; Munch et al., 2006, 2008; Battu et al., 2008; Coulson and Fisher, 2009). Munch et al. (2006) suggest that homeowners may face greater difficulties in moving home than renters but they are more willing to accept local jobs at a lower reservation wage to avoid unemployment. Indeed, homeowners may be prepared to accept the inherent disadvantage associated with homeownership (i.e. being less mobile than renters) and they try to find and keep jobs even if working conditions are unfavourable.

Methodologically, micro-level studies are better placed to cope with the endogeneity of homeownership either through a multi-spell approach (e.g. Van Leuvensteijn and Koning, 2004; Munch et al., 2006, 2008; Battu et al., 2008) or through an instrumental variable approach (e.g. Coulson and Fisher, 2009). Endogeneity of homeownership status may be a problem for two reasons. Firstly, one's homeownership status is correlated with numerous individual and household characteristics which affect preference and the ability to own. If some of these

characteristics are also correlated with unemployment status but not controlled for in the OLS estimation, the estimated coefficient of homeownership is biased and inconsistent (i.e. omitted variable bias). Secondly, one's unemployment status may affect the likelihood of one's homeownership (possibility of reverse causation). Labour income is a significant determinant of homeownership since a stable income stream is a key to making mortgage payments. As unemployment means zero labour income, one might expect that the unemployed have a lower probability of becoming homeowners than the employed. Therefore, both directions of causality could result in a positive correlation between probabilities of becoming a homeowner and being unemployed and the OLS estimation cannot tell us the direction of causality.

The present analysis differs from previous micro-level studies in the UK in terms of methodology. Unlike previous UK studies, it employs an instrumental variable (IV) approach, using the local aggregate homeownership rate and the parents' homeownership status when the children were young as instrumental variables (henceforth instruments). The local homeownership rate is frequently used as an instrument for an individual's homeownership status (e.g. Van Leuvensteijn and Koning, 2004; Dujardin and Goffette-Nagot, 2008; Munch et al., 2008). It can be understood as a comprehensive indicator of how accessible homeownership is in a specific local area. The parents' homeownership status is shown to be a strong predictor of their offspring's homeownership status in the related literature and this is likely because wealth and knowledge of real estate transactions are transferred between generations (Munro, 1989; Hamnett et al., 1991; Deutsch, 1997; Haurin and Morrow-Jones, 2006; Hilber and Liu, 2008). The identifying

assumption is that when various time-varying characteristics and individual fixed effects are controlled for, the parents' past homeownership status is uncorrelated with the children's current unemployment status. In practice, it is used as an instrument in Munch et al. (2008) to explore the causal effect of homeownership on unemployment in Denmark and shows a strong and positive correlation with their children's homeownership status.

No single UK micro-dataset can provide both instruments and hence the two variables need to be obtained from two different datasets. The empirical analysis will also be carried out separately for each instrument. The local homeownership rate is obtainable from the British Household Panel Survey (BHPS) as one of the special-licensed versions of the BHPS provides information on the respondents' local districts of residence. Based on this information, the proportion of homeowners and private renters among the survey respondents are computed for 278 districts of Great Britain. The parents' homeownership status is derived from the British Cohort Study (BCS) 1970 which follows a group of people (cohort members) born in a certain week of 1970 throughout their lifetime. Therefore, the information on the homeownership status of the cohort members' parents can be traced back to when the members were 5 years old. In addition to the instruments, both datasets also contain information on personal and household-level characteristics which help control for variables that are correlated with both unemployment and homeownership status.

A model specification states that one's unemployment status is a function of homeownership status along with other control variables. The empirical section

first estimates the relationship by maximum-likelihood logit estimation. The regression results indicate that the probability of being unemployed is smaller for homeowners than for private renters, contrary to Oswald's hypothesis. However, this finding may only be due to spurious correlation through an individual's innate ability which reduces the likelihood of unemployment but increases that of homeownership. To control for unobserved time-invariant heterogeneity, including innate ability, a fixed effects (FE) model is estimated using the panel structure of the BHPS and the difference in the probability of unemployment between homeowners and private tenants disappears. This result implies that one's homeownership is, indeed, correlated with unobservable characteristics and these are also correlated with unemployment status. However, an FE model cannot account for reverse causality and therefore an FE IV model is estimated with the instrument being the local homeownership rate. The result shows that there is no causal effect of homeownership status on unemployment status. When the model is estimated by 2SLS, using parents' homeownership status as an instrument with the BCS 1970, homeownership is shown again to have no impact on unemployment status.

Overall, this chapter contributes to the gathering of evidence on the causal relationship between homeownership and unemployment. Its key contribution is methodological in that it uses two different types of instruments which have not yet been used with UK data. Both of the instruments are correlated with an individual's homeownership status but have different rationales as an instrument so that the robustness of the empirical results can be double-checked. Furthermore, taking advantage of the panel structure of the BHPS, this chapter runs an FE IV

model which can be more effective in identifying the causality of interest than an IV approach alone.

The structure of the chapter is as follows. Section 2.2 reviews previous papers related to this study and discusses their results to highlight a gap to be filled in the literature. Section 2.3 presents theoretical reasoning and discusses the main testable proposition. Section 2.4 discusses the econometric problems and solutions in detail and introduces the main datasets and explains their advantages for this study. Section 2.5 presents and discusses empirical findings. The last section draws conclusions.

2.2. Literature review

The debate on the relationship between homeownership and unemployment originated from the puzzling phenomenon of high unemployment rates which have been persistent in Western economies since the 1960s. Labour economists led the research at an early stage and paid attention to the *macro-level* relationship between homeownership and unemployment. Oswald (1996) was the first to notice a positive correlation between homeownership and unemployment rates in most advanced countries and suggests that this correlation might be causal as homeowners are less mobile than private renters and hence they would be more vulnerable to negative demand shocks in the local labour market.

Early empirical studies which test the so-called Oswald hypothesis typically use country- or regional-level data. Oswald (1996) estimates a bivariate model of homeownership and unemployment rates using the panel and cross-sectional data for OECD countries and their sub-regions and finds that a 10% rise in the homeownership rate is associated with a 1 to 2% increase in the unemployment rate. Nickell (1998) uses the same data but controls for many more variables in addition to the homeownership rate so that the possibility of omitted variable bias is substantially reduced. The empirical results are still consistent with those of Oswald (1996). Pehkonen (1999) pays sole attention to homeownership as a main cause of unemployment by examining the relationship between the two variables across 13 labour districts in Finland. With no other explanatory variable but the homeownership rate, he concludes that the unemployment rate rises by 1.4 to 2.5% for a 10% increase in the homeownership rate. Green and Hendershott (2001) employ a more sophisticated method using the 1970 and 1990 state-level

homeownership and unemployment rates of the US. Specifically, they use a difference-in-difference estimator to control for state fixed effects, which helps to reduce omitted variable bias. In addition, the detailed grouping of sample by age helps identify positive relationships between the two rates more clearly. Accordingly, for the middle-aged group (34 to 65), the homeownership and unemployment rates are most likely to be positively correlated. Barrios Garcia and Rodriguez Hernandez (2004) address the endogeneity of homeownership in a more explicit manner with Spanish regional data. To deal with the simultaneity of homeownership and unemployment, they rely on a simultaneous equations approach. In contrast to the previous macro-level studies, they report a negative relationship that, at the mean value, implies that a 10% increase in the homeownership rate is associated with a fall of 2.2% in the unemployment rate across regions in Spain.

Despite the methodological improvement over time, a fundamental criticism of macro-level studies is that they test using *macro-data* whereas Oswald's hypothesis describes *individual* homeowners' behaviours and labour market outcomes. It should be questionable whether the macro-level findings are truly the result of the underlying behaviours of individual homeowners. In response to the unjustifiable methodology of macro-level studies, more recent papers tend to depend on micro-data and use more convincing methods to mitigate the endogeneity of homeownership. Van Leuvensteijn and Koning (2004) test the relationship between homeownership and the length of time spent unemployed using data from the Netherlands. To cope with the simultaneity between the two variables, they use simultaneous models of individual job status change and

probability of becoming a homeowner which is instrumented by the regional homeownership rate. For the omitted variable bias in the context of a duration model, they estimate non-parametric models adopting the mass-point methodology introduced by Heckman and Singer (1984). They conclude that residential tenure decisions are driven by job commitment rather than the reverse but that homeowners are less likely to be unemployed than renters.

Munch et al. (2006) examine the same topic using Danish data. Based on search theory, they identify two conflicting factors which jointly determine the relationship between homeownership and unemployment: (1) constrained mobility of homeowners due to high property transaction costs and (2) their willingness to accept locally available jobs at a lower reservation wage to avoid moving costs. Therefore, the overall effects of homeownership may be theoretically ambiguous. Adopting a similar empirical approach to Van Leuvensteijn and Koning (2004), they find that homeowners are likely to find jobs in their local areas rather than in other regions and hence the net impact of homeownership on the period of unemployment is negative. The implication of their findings is that homeowners try to find jobs in other ways than moving to other regions. This may be an expected outcome as homeowners expect low residential mobility and hence they are ready to accept this inherent disadvantage when searching for jobs.

Battu et al. (2008) look at the UK context using the BHPS, also employing a multi-spell approach. Their findings indicate the importance of the initial employment status: homeowners in employment are less likely to transit into jobs

in the distant labour markets while unemployed renters in public housing are associated with a lower probability of finding a job in non-local labour markets. Their results indicate no evidence that homeownership increases the period of unemployment. Furthermore, recent studies show that homeowners are not necessarily immobile when they move for job-related reasons (Hilber and Lyytikäinen, 2012). One defining characteristic of homeowners (relative to renters) is that they face very high relocation costs and one important component of relocation costs is the stamp duty. Hilber and Lyytikäinen (2012), using a Regression Discontinuity Design, find that an increase in the UK stamp duty land tax lowers homeowners' propensities to move substantially but only when they move for housing-related reasons, not when they move for job-related reasons.

The micro-data studies often use a multi-spell approach to deal with the endogeneity problem associated with homeownership. This approach looks at whether and the extent to which homeownership status affects unemployment spells. Its advantage is that it can identify causality without relying on instruments which are normally difficult to find. However, it is not the only solution to the endogeneity of homeownership and its shortcoming is that it is available only from panel data. Furthermore, the length of the unemployment spell is not the only indicator of unfavourable labour market outcomes. The comparison of the probability of being unemployed at a point in time for both homeowners and renters is simpler to test and easier to understand. Coulson and Fisher (2009) compare the probability of being unemployed by residential tenure in the US and rather than using a multi-spell approach, they depend on an IV approach with the state marginal tax rate, percentage of households living in multi-family properties

and sex of children in the households being a set of instruments. They do not find any evidence that homeowners are more likely to be unemployed.

This chapter explores the UK case using an IV approach but does not employ those instruments adopted by Coulson and Fisher (2009) for the following reasons. Firstly, the state-level marginal tax rate is not available in the UK, a centralised country that does not have regional income taxes. Secondly, the sex combination of children is a potentially valid instrument as its determination is exogenous. However, its relevance is challenged empirically, at least in the UK, as it is found to be uncorrelated with the probability of parents being owner-occupiers in the UK BCS 1970 data. It may be driven by the fact that in the UK housing space is much scarcer than in the US and hence not mixing siblings of the opposite sex is more of a luxury in the UK. Lastly, the percentage of households in multi-family properties in the local housing market is highly relevant to one's homeownership status in the UK, too, but it has a high correlation with one's likelihood of being unemployed as well, according to the BHPS. This is thought to be because most multi-family properties are in public housing in the UK and its allocation is closely related to one's employment status and income. Therefore, the public housing tends to be occupied by those who are more likely to be unemployed and an agglomeration of unemployed or less employable people may lower the employability of residents in the neighbourhoods through spill-over or peer effects. This chapter, instead, employs two different instruments: the aggregate homeownership rate and parental homeownership status, the validity and relevance of which will be discussed in the empirical strategy section below.

2.3. Theoretical background

The early macro-level studies suggest a positive relationship between homeownership and unemployment rates without providing a concrete theoretical rationale but rather relying on a loose relationship between geographical mobility and job search efficiency. The micro-studies reviewed above rely on job search theory. Coulson and Fisher (2002) argue that the key insight of their model concerning employment outcomes by tenure type is that renters can enjoy a higher job-matching rate thanks to their higher residential mobility. In Coulson and Fisher (2002), the probability of unemployment for homeowners in the steady state is given as follows:

$$p_u^h = \frac{d}{d + p_l} \tag{1}$$

Where p_u^h is the probability of a homeowner being unemployed, p_l is the probability of match between job seeker and employer in the local labour market and *d* is the probability of a worker being unemployed.

The mechanism is simple: if the matching does not occur at all during a certain period, p_l is zero and hence the unemployment rate (p_u^h) is 1. If p_l is 1 (perfect matching), p_u^h becomes close to zero and homeowners are less likely to be unemployed. The corresponding equation for renters is given as:

$$p_u^r = \frac{d}{d + p_l + p_n} \quad (2)$$

Where p_u^r is the probability of a renter being unemployed and p_n is the probability of match between job seeker and employer in the national labour market.

p_n is added to the denominator of *eq*. (1) reflecting the fact that renters can move for jobs elsewhere and hence their matching probability is the sum of probabilities in both local and national markets. As any probability should range from 0 to 1, p_n is equal to or greater than zero and hence $p_u^h \geq p_u^r$. That is, homeowners are more likely to be unemployed at any point in time.

Munch et al. (2006), however, argue that homeowners may not necessarily be more likely to be unemployed despite their lower residential mobility. This is because they may be more willing to accept available jobs locally at a lower reservation wage than renters arguably because housing transaction costs are much higher for owner-occupiers than renters and so, they may be better off staying put and accepting a lower wage. The implication of their findings for the equations of Coulson and Fisher (2002) is that the matching rate in the *local* labour market is not necessarily the same for both tenure types but it can be higher for homeowners (i.e. $p_l^h > p_l^r$). Therefore, the question of whose unemployment probability is higher depends on the relative size of p_l^h and $p_l^r + p_n$, which can be figured out only empirically.

The discussion above suggests that any factors causing $p_l^h > p_l^r$ can boost the employability of homeowners. There could be various ways for homeowners to cope with their limited job opportunities and hence raise the matching rate in the local job market. Firstly, they may carry out a more rigorous search within their own local areas than renters to raise the matching rate. Rouwendal and Nijkamp (2010) argue that the higher housing expenses involved in homeownership may cause homeowners to search jobs more intensely and hence lower the probability of unemployment or shorten the length of unemployment spell. Secondly, homeowners may have better access to social networks through which they can find better job information for the local labour market, as Coulson and Fisher (2002) point out. Lastly, homeowners may try to lower the likelihood of dismissal while in employment (i.e. $d^r > d^h$) in expectation of limited re-employment opportunities when becoming unemployed. If all of the arguments above hold true, it may even be possible that homeowners are less unemployed, as found in Van Leuvensteijn and Koning (2004). To sum up, theory cannot give a conclusive answer as to the direction of the causal effect of homeownership on unemployment: the direction of the effect is ambiguous and therefore can only be determined empirically.

2.4. Empirical strategy and data

2.4.1. Maximum likelihood logistic model

A basic model specification is a simple binary response model where the dependent variable is a dummy variable taking 1 if a respondent is in unemployment and zero otherwise. The main explanatory variable is also a binary variable referring to the tenure choice between home-owning and private renting. Then, a mathematical specification is:

$$Pr(y_i = 1 | x_i, e_i, D_i) = \frac{exp(\alpha_0 + x_i\beta + D_i\delta + \varepsilon_i)}{1 + exp(\alpha_0 + x_i\beta + D_i\delta + \varepsilon_i)} \quad (3)$$

Where y_i is a dummy for unemployment status, x_i is a dummy for homeowner, D_i is a vector of other controls and ε_i is an error term.

To reduce omitted variable bias to the highest extent possible, in addition to the homeownership dummy, the specification also includes a long list of control variables that may be correlated with both homeownership and unemployment status. The list includes various personal and household characteristics such as age, sex, marital status, number of children and educational level.

The division of the renter group into private and public tenants is important in this chapter as they are highly distinguished from each other in terms of several characteristics, especially, residential mobility. Using the Survey of English Housing, Hughes and McCormick (2000) find that the migration rate (the number of households per 1000 households that moved *between regions* of England) is the highest for private tenants with 4.66% while it is 0.57%, 0.67% and 1.92%

respectively for homeowners, local authority tenants and tenants of social landlord between 1993 and 1998. For all moves, the difference in the mobility is even greater. A move rate (the number of households per 1000 which moved *any distance*) is highest for private renters with 38.75% and lowest for homeowners as 4.23%. Local authority and social tenants are in the middle of the mobility spectrum with rates of 10.94% and 18.36%, respectively. Since the difference in residential mobility is starkest between private renters and homeowners (social renters are more similar to homeowners in terms of their mobility), it seems appropriate to concentrate the analysis that follows on the response of these two groups to external shocks to the local labour market.

2.4.2. Panel-data models

Though it is feasible to control for a large number of household specific variables alongside homeownership, some important characteristics (such as innate ability) may not be observable. Hence, a logit model without fixed effects may lead to biased estimates. In contrast, a panel-data fixed effects model is effective in dealing with unobservable time-invariant characteristics of individuals or households. A specific unobservable variable which deserves attention in the context of homeownership and unemployment is the innate or natural ability of individuals. Since innate ability can affect the likelihood of homeownership positively and that of unemployment negatively, its omission in the logit regression results in a downward bias. Some proxy measures have been suggested in the literature, such as exam or IQ scores but none of them are perfect measures and, in any case, they are not available from the BHPS. Instead, using the panel structure of the BHPS, a fixed effects (FE) model is estimated as a partial solution

to unobserved individual characteristics including innate ability. The solution is only partial since fixed effects only control for time-invariant but not time-varying unobservable characteristics. Still, estimating an FE model is arguably a significant improvement over standard OLS or logit models.

Other than the FE model, another well-known panel model is the random effects (RE) model which treats unobserved characteristics as random and leaves them in a composite error term. Its estimator is the most efficient as it takes advantage of both within- and between-variations but may be inconsistent if the orthogonality condition of $cov(x_{it}, u_i) = 0$ is not satisfied (where u_i refers to unobserved individual characteristics). In this case, homeownership status and innate ability are likely to be correlated and therefore an RE model is not expected to deliver a consistent estimator. However, it is worth looking at its outcomes as it can be used to ascertain the degree of the bias caused by the time-constant unobservable characteristics through the comparison of the results from the corresponding FE model.

As the dependent variable is a binary variable, it might seem preferable to use an FE logit model rather than an FE linear probability model (LPM). The latter is not ideal for dealing with a binary outcome variable as it allows for implausible estimated probabilities (greater than 1 or smaller than 0) for the outcome variable. Due to the complicated nature of the transformation process required for the FE logit model, those households that exhibit no variation in value of the dependent variable over time (i.e. those who were never unemployed or never employed)

have to be excluded from the estimation.[1] Because of the persistency in the employment status over time, a substantial number of households are lost when estimating an FE logit model. This issue may suggest employing, despite its shortcomings, an FE linear probability model. Results are therefore also reported for an FE linear probability model in order to check whether the findings are sensitive to the choice of estimator.

2.4.3. Identification strategy

The panel-data fixed effects models are a partial solution to deal with omitted variable bias since they do not have the power to deal with unobserved *time-variant* individual characteristics. One example of time-variant characteristics is an unobserved preference for mobility which may not be constant over time and can be affected by life-cycle status, employment status or working conditions. More importantly, bias from mutual causality between homeownership and employment status cannot be dealt with by panel fixed effects models. The income is, needless to say, one of the most significant determinants of homeownership. As a main source of stable income and a crucial condition for mortgage qualification, employment becomes one of the most influential determinants of homeownership. Therefore, a negative correlation between unemployment status and homeownership status is represented by not only the causality from the latter to the former but also the reverse. To identify only the causality this chapter focuses, an IV approach is needed. In this chapter, two types of instruments are considered: (1) aggregate homeownership rates at the levels of local authority district and county and (2) parental homeownership status.

[1] See Frees (2004) for detailed discussion on sample loss in the FE logit model.

2.4.3.1. Aggregate homeownership rate as an instrument

The aggregate homeownership rate at a specified geographical level has been frequently used as an instrument when the causal effects of homeownership on various social phenomena need to be identified. The reason for its popularity as an instrument is that it can be regarded as a comprehensive indicator of how accessible homeownership is in a specific local area. The rate reflects the characteristics related to the propensity to own (e.g. income, wealth and mobility) of average households in that area. Also, it is closely related to conditions in the local housing market which affect individuals' tenure decisions (e.g. availability of housing stock suitable for ownership). Munch et al. (2008) argue that the regional homeownership rate affects individual homeownership status through a supply effect. Therefore, it may be a better predictor of an individual's tenure choice than any single individual determinant of homeownership. The special-licensed version of BHPS data used in this chapter provides information on three levels of geographical units of survey sample – region, county and local authority district for each wave.[2] Aggregate homeownership rates at county and district levels will be used as they are likely to be more strongly related with homeownership statuses of individuals than regional homeownership rates are. For each wave of the BHPS, the local homeownership rate would be computed as the ratio of owner-occupying households to total households for each of 57 counties and 278 districts as in 1991 when the BHPS began.

[2] The title of the data is 'British Household Panel Survey, Waves 1-18, 1991-2009: Conditional Access for Non-UK Users, Local Authority District Codes'. The details of the dataset are found on the website of 'Economic and Social Data Service' (www.esds.ac.uk). The dataset used for this chapter is an earlier version of this dataset and hence the local district codes are available from 1991 to 2006.

2.4.3.2. Parental homeownership status as an instrument

A second instrument to be employed is the parental homeownership status which has also been employed by Munch et al. (2008). In the literature, children whose parents were homeowners have a greater propensity to owner-occupy when they are grown-ups (Henretta, 1984; Boehm and Schlottman, 1999; Mulder and Smits, 1999). The inter-generational correlation of homeownership seems to have two sources. Firstly, there may be a transmission of knowledge of home-buying from parents to children such as knowledge on housing market searching (e.g. Farley, 1996; Hirad and Zorn, 2002), use of credit markets (e.g. Courchane, Surette and Zorn, 2004) and specific local characteristics of target markets such as neighbourhood quality (e.g. Hilber, 2005), which are demonstrated to play significant roles in the home-buying process. One source from which to acquire such knowledge, free from any third-party interest, are parents who have experience regarding housing transactions. Haurin and Morrow-Jones (2006) show that additional market knowledge from parents does matter for children's homeownership decisions in the US based on their own survey data.

Secondly, there may be an inter-generational transfer of wealth which not only makes a direct financial contribution but also eases the qualification test for mortgage acquisition. Munro (1989) and Hamnett et al. (1991) confirm the importance of intra-family transfer of wealth as the source of homeownership in the UK as cited in Deutsch (1997) who also shows the same result using Austrian data. Engelhardt and Mayer (1994) find that about 20% of first-time homebuyers obtain financial help from their family or kin worth an average of about 50% of the down payment. Charles and Hurst (2002) suggest that Afro-Americans have a

higher decline rate of mortgage applications than White because they have a lower level of financial assistance from family. Hilber and Liu (2008) demonstrate that the previously unexplained gap in homeownership propensity between Black and White households in the US, of about 6.5%, disappears entirely when a household's own wealth, parental wealth and macro-location preference are jointly controlled for. In the UK, 38% of first-time home buyers under age 30 had assistance with the down-payment in 2006 (Tatch, 2007).

In order to be valid as an instrument, parental homeownership cannot be correlated with the error term. In practical terms, this means that parental homeownership ought to be uncorrelated with unobserved individual characteristics of the offspring that may affect the offspring's employment status. It has been reported in the literature that parental homeownership status exerts many positive effects on child outcomes such as lower level of school drop-outs, lower pregnancy rates among teenagers (Green and White, 1997; Aaronson, 2000), higher exam results and fewer behavioural problems (Haurin et al, 2001). The fact that better educational results are closely related to employment outcomes suggests that parental homeownership status cannot be regarded as completely irrelevant to employment outcomes of the children. Therefore, for parental homeownership status to be a valid instrument, one's educational characteristics need to be included in the model.

2.4.4. Data

2.4.4.1. British Household Panel Survey

One of the two datasets for the empirical analysis is the British Household Panel Survey (BHPS). It is a panel survey which has been carried out annually. The data from 1991 to 2006 (16 waves) are used in this chapter. The survey follows over time an initial sample of about 10,000 individuals who are over 16 year old and these individuals belong to 5,500 representative households in Great Britain. The sample of households is occasionally boosted to compensate for households that exit the panel. The survey topics are various and the most relevant types of information for this chapter are basic demographic characteristics, labour market status and job characteristics, residential tenure, accommodation characteristics, education and qualifications. It is well-known that panel data have a number of benefits unavailable in either cross-sectional or time-series data alone. In particular, the panel structure of the data, by employing fixed effects, is often used to control for unobserved individual heterogeneity and this advantage is of great help to the present research topic.

The total sample size is 209,714 individuals over 16 waves, out of which, 164,225 are of working age (16 to 64 year old for men and 16 to 59 for women). When public renters are excluded, 132,087 observations remain and the respondents whose districts (and hence the local homeownership rates) are known are 119,179. After the households with any missing values for the variables entering the specification are excluded, the final sample size for the regressions is 113,839. Table 2.1 summarises the basic statistics for the variables from the BHPS.

Table 2.1. Summary statistics for the BHPS variables

Variable	Obs.	Mean	Std. Dev.	Min	Max
Respondent was unemployed	113839	0.034	0.181	0	1
Respondent was homeowner (private renter exc.)	113839	0.892	0.310	0	1
Age	113839	38.133	12.704	16	64
Age squared	113839	1615.491	991.633	256	4096
Sex (Female exc.)	113839	0.493	0.500	0	1
HH head (Non-head exc.)	113839	0.474	0.499	0	1
Qualification (First or higher degree exc.)					
Other higher qualification	113839	0.264	0.441	0	1
A level	113839	0.167	0.373	0	1
GCSE	113839	0.229	0.420	0	1
Other qualification	113839	0.057	0.232	0	1
No qualification	113839	0.128	0.334	0	1
Marital status (Married exc.)					
Separated	113839	0.020	0.139	0	1
Divorced	113839	0.075	0.263	0	1
Widowed	113839	0.012	0.111	0	1
Never married	113839	0.318	0.466	0	1
No. of dependent children in the HH	113839	0.606	0.953	0	8
No. of HH members	113839	3.101	1.295	1	14
District homeownership rate	113839	0.702	0.144	0	1
District unemployment rate	113839	0.038	0.035	0	1
District employment rate	113839	0.581	0.103	0	1
County homeownership rate	113839	0.695	0.081	0	1
County unemployment rate	113839	0.039	0.019	0	0.143
County employment rate	113839	0.578	0.062	0	0.889
Region (Inner London exc.)					
Outer London	113839	0.024	0.154	0	1
Rest of South East	113839	0.051	0.219	0	1
South West	113839	0.169	0.375	0	1
East Anglia	113839	0.078	0.268	0	1
East Midlands	113839	0.035	0.184	0	1
West Midlands conurbation	113839	0.071	0.256	0	1
Rest of West Midlands	113839	0.027	0.162	0	1
Greater Manchester	113839	0.044	0.206	0	1
Merseyside	113839	0.033	0.178	0	1
Rest of North West	113839	0.018	0.132	0	1
South Yorkshire	113839	0.040	0.195	0	1
West Yorkshire	113839	0.022	0.146	0	1

Table 2.1. Summary statistics for the BHPS variables (cont.)

Variable	Obs.	Mean	Std. Dev.	Min	Max
Rest of Yorkshire & Humberside	113839	0.026	0.160	0	1
Tyne & Wear	113839	0.029	0.167	0	1
Rest of North	113839	0.017	0.130	0	1
Wales	113839	0.032	0.177	0	1
Scotland	113839	0.133	0.340	0	1
Year (2006 exc.)					
1991	113839	0.055	0.228	0	1
1992	113839	0.053	0.224	0	1
1993	113839	0.051	0.219	0	1
1994	113839	0.051	0.221	0	1
1995	113839	0.050	0.218	0	1
1996	113839	0.052	0.222	0	1
1997	113839	0.056	0.230	0	1
1998	113839	0.055	0.227	0	1
1999	113839	0.079	0.270	0	1
2000	113839	0.080	0.271	0	1
2001	113839	0.078	0.269	0	1
2002	113839	0.070	0.256	0	1
2003	113839	0.068	0.252	0	1
2004	113839	0.066	0.249	0	1
2005	113839	0.068	0.251	0	1

2.4.4.2. British Cohort Study 1970

The second dataset is the British Cohort Study (BCS) which has been carried out since 1970. The study traces a group of about 17,200 individuals (cohort members) who were born in Great Britain in one particular week during 1970. It also has panel characteristics as the individuals are repeatedly surveyed over time but it is not as regular as the BHPS. Since the first survey, six major follow-up surveys have been carried out at the ages of 5, 10, 16, 26, 30, 34 and 38, with the most recent one done in 2008. Initially, the survey topics were mainly confined to medical conditions of new born babies. As the individual cohort members grew up, the topics have also been expanded to social, educational and occupational characteristics. A particular benefit from the BCS 1970 is that it allows for access to information on parental homeownership status when the cohort members were young. For the empirical analysis, the 2004 survey will be employed because a relatively permanent tenure type is likely to be determined by the age of 34.[3] At earlier ages (26 or 30), most individuals may still be in transition to their preferred tenure type.

The original cohort members are 17,196 but the number of members surveyed is reduced substantially to 9,665 by 2004. Of these, 7,828 are homeowners or private renters and public renters and other types of renters are excluded at this stage. Among the homeowners and private renters, the information on parents' homeownership status when they were 5 years old is available only for 6,457. The final number of the regression sample is 6,118 after observations with missing values for regression variables are excluded. See Table 2.2 for summary statistics.

[3] At the time of empirical experiment, the 2004 survey was the most recent one.

Table 2.2. Summary statistics for the BCS 1970 variables

Variable	Obs.	Mean	Std. Dev.	Min	Max
Respondent is unemployed	6118	0.014	0.116	0	1
Respondent is homeowner (private renter exc.)	6118	0.918	0.274	0	1
Age	6118	33.882	0.356	33	35
Age squared	6118	1148.083	23.920	1089	1225
Sex (Female exc.)	6118	0.477	0.500	0	1
Qualification (Higher degree exc.)					
First degree	6118	0.142	0.349	0	1
A level	6118	0.334	0.472	0	1
GCSE	6118	0.098	0.297	0	1
CSE	6118	0.291	0.454	0	1
No qualification	6118	0.072	0.259	0	1
Marital status (Married exc.)					
Cohabiting	6118	0.205	0.404	0	1
Single (never married)	6118	0.144	0.351	0	1
Separated, divorced, widowed	6118	0.049	0.216	0	1
No. of children	6118	1.089	1.052	0	7
No. of HH members	6118	2.062	1.239	0	8
Father's occupational class when respondent was 5 year old. (Professional occupations exc.)					
Managerial and low professional occupations	6118	0.229	0.420	0	1
Non-manual skilled occupations	6118	0.100	0.299	0	1
Manual skilled occupations	6118	0.442	0.497	0	1
Semi-skilled occupations	6118	0.115	0.320	0	1
Unskilled occupations	6118	0.032	0.177	0	1
Father was homeowner when respondent was 5 year old.	6118	0.659	0.474	0	1
Region (Scotland exc.)					
North East	6118	0.048	0.213	0	1
North West	6118	0.117	0.321	0	1
Yorkshire & Humberside	6118	0.099	0.298	0	1
East Midlands	6118	0.073	0.261	0	1
West Midlands	6118	0.097	0.296	0	1
East of England	6118	0.110	0.313	0	1
London	6118	0.088	0.283	0	1
South East	6118	0.146	0.353	0	1
South West	6118	0.088	0.283	0	1
Wales	6118	0.056	0.229	0	1

2.5. Empirical results

2.5.1. Pooled logit regressions with BHPS

The results of logit regression with a pooled sample from the BHPS are discussed first. Though a logit model specification is hardly capable of dealing with the omitted variable bias and the endogeneity of the homeownership status, it allows us to ascertain the biases caused by it when comparing the findings to those of panel fixed effects estimates or IV estimation. Table 2.3 shows the results of binary logit regressions of an individual's unemployment on homeownership status and other control variables. The control variables are gradually added from left to right columns. In column (1), only the dummy for homeownership status is controlled for along with regional and year dummies. The main finding is that the estimated coefficient of homeownership status is negative implying that homeowners are less likely to be unemployed than private renters. However, it is likely biased due to incomplete control of explanatory variables. In columns (2) and (3), major individual demographic characteristics and household characteristics are added respectively. The inclusion of more control variables makes the estimated coefficients of the homeownership variable smaller, which implies that the omitted variable bias is reduced. However, even after a number of variables arguably correlated with both unemployment and homeownership statuses are controlled for, the differences in the probability of unemployment remains significant. Some unobservable heterogeneity may be responsible for the difference. This problem is addressed by employing a panel fixed effects model, discussed in the next section.

Table 2.3. Logit estimation of unemployment

(Dependent variable: unemployment status (1 if unemployed, otherwise 0))

	(1) HO status only	(2) Personal characteristics	(3) HH characteristics
Homeowner (Private renter exc.)	-0.836***	-0.741***	-0.590***
	(0.042)	(0.045)	(0.045)
Age		-0.085***	-0.013
		(0.008)	(0.010)
Age squared		0.001***	0.000*
		(0.000)	(0.000)
Male (Female exc.)		0.727***	0.633***
		(0.044)	(0.040)
Head of HH (Non-head exc.)		-0.385***	-0.326***
		(0.045)	(0.044)
Qualification (First or higher degree exc.)			
Other higher qual.		-0.128**	-0.089
		(0.058)	(0.059)
A level		-0.319***	-0.293***
		(0.065)	(0.066)
GCE O level		0.005	0.051
		(0.060)	(0.061)
Other qualification		0.360***	0.412***
		(0.074)	(0.075)
No qualification		0.658***	0.705***
		(0.062)	(0.063)
Marital status (Married exc.)			
Separated			1.212***
			(0.100)
Divorced			0.875***
			(0.064)
Widowed			0.560***
			(0.164)
Never married			1.018***
			(0.059)
No. of dependent children in HH			-0.073**
			(0.029)
No. of HH members			0.078***
			(0.016)
Region	Yes	Yes	Yes
Year	Yes	Yes	Yes
Constant	-2.953***	-1.429***	-3.961***
	(0.087)	(0.172)	(0.231)
Pseudo. R-squared	0.0224	0.0530	0.0685
No. of obs.	113,839	113,839	113,839

Note: ***, **,* refer to the significance levels at 1, 5 and 10% respectively. Robust standard errors are given in parenthesis. The estimated coefficients for the regional and year dummies are omitted from the table.

2.5.2. Panel models

Table 2.4 presents a few panel model regression results and the estimation of the Random Effects (RE) model is presented first in column (1). The RE estimate of the coefficient of the homeowner dummy is negative and significant at the 1% level, which is the same result as the pooled logit estimation in column (3) of Table 2.3. However, it is likely to be biased due to the presence of unobservable personal characteristics which affect both the unemployment and the homeownership statuses. The ρ-value associated with the RE model regression is 0.49, which means that the unobserved individual characteristics account for about 49% of variation in the composite errors. A fixed effects model regression is needed to control for unobservable personal characteristics.

In column (2) of Table 2.4, the logit FE model estimation result is given. The estimated coefficient of the homeownership variable turns positive but is not statistically significant. This suggests that there may be no difference in the probability of unemployment between homeowners and private renters. Though not as efficient as the RE model, the virtue of the FE estimator is that it is consistent even when unobserved time-constant characteristics are present. Hence, the estimates from the RE model are considered consistent only when they are not significantly different from the corresponding FE estimates. The estimated coefficients of the homeownership variable look different at a glance between the FE and RE models. The difference is also confirmed by a Hausman specification test. Hence, the RE estimate of the homeownership variable in column (1) is formally demonstrated to be biased.

Table 2.4. Fixed and Random effects logit estimation of unemployment
(Dependent variable: unemployment status (1 if unemployed, otherwise 0))

	(1) RE logit	(2) FE logit	(3) FE LPM
Homeowner	-0.508***	-0.107	-0.000
(Private renter exc.)			
	(0.063)	(0.098)	(0.003)
Age	-0.010	-0.170**	-0.006***
	(0.014)	(0.080)	(0.002)
Age squared	0.000	0.000	0.000***
	(0.000)	(0.000)	(0.000)
Male (Female exc.)	0.730***		
	(0.061)		
Head of HH (Non-head exc.)	-0.381***	-0.160*	-0.006**
	(0.062)	(0.092)	(0.002)
Qualification			
(First or higher degree exc.)			
Other higher qual.	-0.304***	-2.543***	-0.062***
	(0.087)	(0.221)	(0.005)
A level	-0.549***	-2.239***	-0.063***
	(0.091)	(0.203)	(0.005)
GCE O level	-0.257***	-3.072***	-0.081***
	(0.089)	(0.231)	(0.005)
Other qualification	0.238**	-2.679***	-0.062***
	(0.119)	(0.317)	(0.008)
No qualification	0.561***	-3.106***	-0.075***
	(0.098)	(0.284)	(0.007)
Marital status (Married exc.)			
Separated	1.315***	0.624***	0.021***
	(0.131)	(0.172)	(0.005)
Divorced	0.860***	0.227	0.005
	(0.095)	(0.154)	(0.004)
Widowed	0.726***	0.313	0.013
	(0.233)	(0.429)	(0.009)
Never married	1.010***	0.447***	0.009***
	(0.083)	(0.141)	(0.003)
No. of children in HH	-0.144***	-0.325***	-0.008***
	(0.037)	(0.055)	(0.001)
No. of workers in HH	0.082***	0.114***	0.005***
	(0.022)	(0.033)	(0.001)
Region	Yes	Yes	Yes
Year	Yes	Yes	Yes
Constant	-5.043***		0.260***
	(0.329)		(0.088)
Adj. R-squared		0.0524	0.0026
No. of obs.	113,839	19,517	113,839

Note: ***, **,* refer to the significance levels at 1, 5 and 10% respectively. Robust standard errors are given in parenthesis. The estimated coefficients for the regional and year dummies are omitted from the table. The sample size for the result shown in column (2) is less than those in the other two columns due to the transformation process required for the logit FE model. The R-squared for LPM FE model in column (3) is the overall R-squared.

Though the FE logit model produces consistent estimates immune from unobservable heterogeneity across individuals, its expenses are not small in terms of efficiency loss due to the requirement for the transformation as discussed earlier. Any respondents who have not experienced any changes in their unemployment status over the entire period of the surveys (16 waves here) would be excluded from the sample to be used for the estimation. The loss is particularly large since employment or unemployment status tends to be persistent over time. Out of 113,839 observations for 18,153 individuals, only 19,517 observations of 2,182 individuals are used for the logit FE estimation in column (2) of Table 2.4. The model does not make use of the data efficiently and hence standard errors are larger than they would be with a larger sample size. As a consequence, one might be concerned that the null hypothesis of no difference in the probability of unemployment between homeowners and private renters is less likely to be rejected.

To check the robustness of the FE logit model result, an FE linear probability model (LPM) is used as a supplementary model in column (3) of Table 2.4. Unlike the FE logit model, it does not require the strict transformation condition and hence there is no loss of sample. It drops an assumption of non-linearly distributed error terms and is not the best model to deal with binary models but it is expected to reveal whether the reduction in sample size affects the sign and significance of the logit FE estimate. Though a direct comparison of the size of estimates between FE logit and LPM is not possible, they do indicate the same finding that there is no correlation between homeownership and unemployment. Therefore, the FE logit does not seem to be affected by the loss of sample.

Overall, it is clearly shown that unobserved characteristics are indeed relevant through the comparison among pooled logit, RE and FE models. When they are properly dealt with by FE models, the negative correlation between homeownership and unemployment disappears. Then, it could be cautiously suggested that the estimated coefficient of the homeownership status may even turn positive if the possibility of reverse causation from unemployment to homeownership is removed. The next section will attempt to discount the concern of reverse causation through 2SLS estimation.

2.5.3. 2SLS estimation with homeownership rate as an instrument

Table 2.5 shows the IV estimation results with the pooled sample of the BHPS. As discussed, the instrument is the aggregate homeownership rates. One of the key requirements for an instrument is that it should be correlated with the dependent variable only through the independent variable to be instrumented but not with the error term (i.e. through omitted variables). In this specific case, however, the homeownership rate of a certain district (the instrument) could be correlated with the unemployment statuses of the individual residents (the dependent variable) of that district through variables other than the homeownership status (the endogenous variable to be instrumented), particularly, through the local unemployment and employment rates for two reasons. Firstly, there are a relatively small number of sample for each district in the BHPS. Secondly, the BHPS picks up the households from a relatively limited number of postcode areas within a certain district. As a result, the BHPS surveyees from the same district tend to share more homogenous neighbourhood characteristics than the entire

households from that district do. Therefore, the local unemployment and employment rates which are calculated based on the small number of relatively homogeneous households are likely to be closely correlated with the employment status of the individual members of those households. If the local homeownership rate is correlated with unemployment and employment rates within a district - which is likely as both of them are computed from the relatively small number of households in each district of the BHPS - the local homeownership rate could also be correlated with the individual's unemployment status through the local unemployment and employment rates. Then, the local homeownership rate would not be a valid instrument if unemployment rates and employment rates remain in the error terms. Therefore, those two rates should be controlled for in the regression specification.

Before the IV results are discussed, the comparable OLS estimation result in column (1) of Table 2.5 is explained first to demonstrate the bias introduced by endogeneity of homeownership. As a binary dependent model combined with an IV approach is not easy to estimate with a logit model, it is estimated by OLS (i.e. linear probability model), which should still be able to show the sign and significance level of the estimated coefficient of interest correctly. The linear probability model with the pooled sample over 16 waves shows again that homeowners are associated with a lower probability of unemployment. The estimated coefficient of the homeownership status has a negative sign and is significant at the 1% level. It is consistent with the results from the pooled logit estimation in column (3) of Table 2.3.

Table 2.5. IV estimation of unemployment with BHPS

(Dependent variable: unemployment status (1 if unemployed, otherwise 0))

	(1) OLS	(2) IV with district HO rate	(3) IV with county HO rate
Homeowner (Private renter exc.)	-0.026***	-0.020	-0.001
	(0.002)	(0.028)	(0.098)
Age	-0.000	-0.001	-0.001
	(0.000)	(0.001)	(0.001)
Age squared	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Male (Female exc.)	0.027***	0.026***	0.025***
	(0.002)	(0.002)	(0.006)
Head of HH (Non-head exc.)	-0.018***	-0.018***	-0.016
	(0.002)	(0.003)	(0.010)
Qualification (First or higher degree exc.)			
Other higher qual.	-0.003**	-0.003**	-0.004
	(0.002)	(0.002)	(0.003)
A level	-0.010***	-0.010***	-0.010***
	(0.002)	(0.002)	(0.002)
GCE O level	0.000	-0.000	-0.001
	(0.002)	(0.002)	(0.005)
Other qualification	0.015***	0.015***	0.014***
	(0.003)	(0.003)	(0.005)
No qualification	0.024***	0.024***	0.025***
	(0.002)	(0.002)	(0.003)
Marital status (Married exc.)			
Separated	0.043***	0.044***	0.047***
	(0.005)	(0.006)	(0.014)
Divorced	0.028***	0.029***	0.030***
	(0.002)	(0.003)	(0.006)
Widowed	0.021***	0.021***	0.020***
	(0.005)	(0.005)	(0.006)
Never married	0.031***	0.032***	0.034***
	(0.002)	(0.004)	(0.011)
No. of children in HH	-0.002***	-0.002***	-0.002*
	(0.001)	(0.001)	(0.001)
No. of HH members	0.003***	0.003***	0.003
	(0.001)	(0.001)	(0.002)
District unemployment rate	0.143***	0.144***	
	(0.022)	(0.022)	
District employment rate	-0.026***	-0.027***	
	(0.006)	(0.008)	
County unemployment rate			-0.059**
			(0.024)
County employment rate			0.123***
			(0.041)
Region	Yes	Yes	Yes
Year	Yes	Yes	Yes
Constant	0.054***	0.038**	0.046
	(0.010)	(0.016)	(0.046)
Adj. R-squared	0.023	0.023	0.020
No. of obs.	113,839	113,839	113,839

Note: ***, **,* refer to the significance levels at 1, 5 and 10% respectively. Robust standard errors are given in parenthesis. The estimated coefficients for the regional and year dummies are omitted from the table. See Table 2.6 for the related first-stage regressions. The endogenous variable is the homeownership status and the instrumental variables are the district homeownership rate for column (2) and county-level homeownership rate for column (3).

A comparable IV estimation result with the instrument being the district-level homeownership rate is presented in column (2) of Table 2.5. The corresponding first-stage regression result is shown in column (1) of Table 2.6. The first-stage regression result shows that the homeownership rate is a highly relevant instrument to the individual homeownership status. Firstly, the homeownership rate variable has a statistically significant effect on the individual homeownership status and it is positive as expected. Roughly speaking, a 1% increase in the local homeownership rate at the district level raises the probability of homeownership by 0.157%. Secondly, the F statistic from the first-stage regression is significant at the 1% level and its p-value is practically zero. Lastly, the Kleibergen-Paap Wald F statistic is 390.6, suggesting that weak identification does not appear to be a problem here. An over-identification test cannot be performed as the equation is exactly identified.

The most noteworthy change is that the estimated coefficient of interest by the IV estimation becomes less negative and is no longer statistically significant implying that homeownership may not causally affect unemployment. Overall, there is no relationship found between one's homeownership and unemployment status when reverse causality is dealt with by IV estimation. It is possible that there may exist other types of local characteristics than the unemployment and employment rates which connect the local homeownership rates and individuals' unemployment statuses. However, it is not feasible to control for an infinite number of aggregate local characteristics and it is not justifiable to do so without concrete theoretical reasons.

Table 2.6. First stage regression of IV estimation with BHPS

(Dependent variable: homeownership status (1 if homeowner, 0 private renter))

	(1) First stage regression for column (2) of Table 2.5	(2) First stage regression for column (3) of Table 2.5
Age	0.014***	0.014***
	(0.001)	(0.001)
Age squared	-0.000***	-0.000***
	(0.000)	(0.000)
Male (Female exc.)	0.059***	0.060***
	(0.003)	(0.003)
Head of HH (Non-head exc.)	-0.100***	-0.101***
	(0.003)	(0.003)
Qualification (First or higher degree exc.)		
Other higher qual.	0.026***	0.027***
	(0.003)	(0.003)
A level	-0.006*	-0.006*
	(0.003)	(0.003)
GCE O level	0.049***	0.050***
	(0.003)	(0.003)
Other qualification	0.033***	0.033***
	(0.004)	(0.004)
No qualification	-0.019***	-0.020***
	(0.003)	(0.003)
Marital status (Married exc.)		
Separated	-0.135***	-0.136***
	(0.009)	(0.009)
Divorced	-0.056***	-0.058***
	(0.004)	(0.004)
Widowed	0.047***	0.046***
	(0.006)	(0.006)
Never married	-0.103***	-0.106***
	(0.003)	(0.003)
No. of children in HH	-0.010***	-0.009***
	(0.001)	(0.001)
No. of workers in HH	0.017***	0.016***
	(0.001)	(0.001)
District unemployment rate	-0.022	
	(0.036)	
District employment rate	0.103***	
	(0.012)	
District homeownership rate	0.157***	
	(0.008)	
County unemployment rate		0.103
		(0.065)
County employment rate		0.183***
		(0.021)
County homeownership rate		0.103***
		(0.017)
Region	Yes	Yes
Year	Yes	Yes
Constant	0.404***	0.391***
	(0.016)	(0.021)
Adj. R-squared	0.122	0.117
Kleibergen-Paap Wald F statistic	390.570	35.806
No. of obs.	113,839	113,839

Note: ***, **,* refer to the significance levels at 1, 5 and 10% respectively. Robust standard errors are given in parenthesis. The estimated coefficients for the regional and year dummies are omitted from the table.

An alternative way to check the robustness of the IV estimation result would be to analyse the homeownership rate from a larger geographical unit, which would mean the local homeownership rate may be less relevant to individual homeownership status but, at the same time, more exogenous. The 278 local districts are grouped into the corresponding 57 counties as of 1991 and the county-level homeownership rate is calculated. It is an empirical question whether the homeownership rate at the county level can still work as a relevant instrument while being exogenous enough. At least, the relevant first stage regression in column (2) of Table 2.6 confirms that it is also highly correlated with an individual's homeownership status. Also, the diagnostic statistics support its relevance: the Kleibergen-Paap Wald F statistic is 35.8. Now the estimated coefficient of the homeownership status increases further but is not statistically different from zero. Overall, these results are strongly suggestive that one's homeownership status does not cause unemployment.

Finally, there may be the possibility that unobservable time-constant characteristics of the individuals are correlated with the local homeownership rate because it is calculated based on a small-sized sample. To control for the unobservable heterogeneity, a FE IV model is estimated and its result is presented in Table 2.7. The deviation of homeownership status from its average over time is instrumented by the deviation of homeownership rates from their average. As the change in the homeownership rate is not great over time, there may be a concern that the deviation of homeownership rates would be fairly small and hence the homeownership rate may not be as highly relevant an instrument as in the case of pooled IV regressions.

Table 2.7. Fixed effects IV estimation of unemployment with BHPS
(Dependent variable: unemployment status (1 if unemployed, otherwise 0))

	(1) IV with district homeownership rate	(2) IV with county-level homeownership rate
Homeowner (Private renter exc.)	0.090	0.102
	(0.070)	(0.134)
Age	-0.006***	-0.007***
	(0.002)	(0.003)
Age squared	0.000***	0.000***
	(0.000)	(0.000)
Head of HH (Non-head exc.)	0.006	0.008
	(0.010)	(0.018)
Qualification (First or higher degree exc.)		
Other higher qual.	-0.061***	-0.060***
	(0.005)	(0.006)
A level	-0.057***	-0.056***
	(0.007)	(0.010)
GCE O level	-0.084***	-0.084***
	(0.006)	(0.007)
Other qualification	-0.062***	-0.061***
	(0.009)	(0.009)
No qualification	-0.076***	-0.075***
	(0.009)	(0.009)
Marital status (Married exc.)		
Separated	0.027***	0.028***
	(0.007)	(0.010)
Divorced	0.006	0.006
	(0.004)	(0.004)
Widowed	0.007	0.006
	(0.009)	(0.012)
Never married	0.018**	0.020
	(0.008)	(0.014)
No. of children in HH	-0.007***	-0.007***
	(0.001)	(0.002)
No. of HH members	0.002	0.002
	(0.003)	(0.005)
District unemployment rate	0.122***	
	(0.023)	
District employment rate	0.002	
	(0.010)	
County unemployment rate		0.023
		(0.046)
County employment rate		-0.008
		(0.019)
Adj. R-squared	.	.
No. of obs.	109,932	109,932

Note: ***, **,* refer to the significance levels at 1, 5 and 10% respectively. Robust standard errors are given in parenthesis. The estimated coefficients for the regional and year dummies are omitted from the table. R^2 are not reported as they are negative. See Table 2.8 for the related first-stage regressions. The endogenous variable is the homeownership status and the instrumental variables are the district homeownership rate for column (1) and county-level homeownership rate for column (2). The number of observations used is 3,907 less than the total sample size (113,839 observations). The reduction in the sample size is due to the singleton groups where only one observation exists.

As can be seen in the related first-stage regression results presented in Table 2.8 below, it turns out that the concern explained above is not serious. In column (1) of Table 2.8 where the instrument is the district-level homeownership rate, the size of the estimates for the homeownership rate is lower than the counterpart in the non-panel first-stage regressions in column (1) of Table 2.6 but it still has a predicted sign and is statistically significant at the 1%-level. In column (2) of Table 2.8 where the instrument is the county-level homeownership rate, the size and the significance are not different from those of the corresponding estimate in column (2) of Table 2.6 (i.e. pooled IV regressions). Again, the Kleibergen-Paap Wald F statistics in Table 2.8 become lower compared to those in Table 2.6 but they are still large enough to support the relevance of the homeownership rates as instruments. A main finding from the second-stage regressions in Table 2.7 is that the estimated coefficients of the homeownership status are now positive but they are still insignificant, independent of whether the district- or county-level homeownership rate is used as an instrument.

Table 2.8. First stage regression of FE IV estimation with BHPS
(Dependent variable: homeownership status (1 if homeowner, 0 private renter))

	(1) First stage regression for column (1) of Table 2.7	(2) First stage regression for column (2) of Table 2.7
Age	0.007**	0.007**
	(0.003)	(0.003)
Age squared	-0.000*	-0.000*
	(0.000)	(0.000)
Head of HH (Non-head exc.)	-0.133***	-0.135***
	(0.007)	(0.007)
Qualification (First or higher degree exc.)		
Other higher qual.	-0.024**	-0.023*
	(0.012)	(0.012)
A level	-0.062***	-0.064***
	(0.013)	(0.013)
GCE O level	0.027**	0.027**
	(0.013)	(0.013)
Other qualification	-0.003	-0.003
	(0.016)	(0.017)
No qualification	-0.001	-0.003
	(0.016)	(0.016)
Marital status (Married exc.)		
Separated	-0.063***	-0.063***
	(0.012)	(0.012)
Divorced	-0.007	-0.007
	(0.010)	(0.010)
Widowed	0.073***	0.073***
	(0.014)	(0.014)
Never married	-0.101***	-0.103***
	(0.008)	(0.009)
No. of children in HH	-0.013***	-0.013***
	(0.002)	(0.002)
No. of workers in HH	0.034***	0.034***
	(0.002)	(0.002)
District unemployment rate	-0.091***	
	(0.034)	
District employment rate	0.064***	
	(0.021)	
District homeownership rate	0.092***	
	(0.016)	
County unemployment rate		-0.101
		(0.065)
County employment rate		0.014
		(0.038)
County homeownership rate		0.103***
		(0.029)
Region	Yes	Yes
Year	Yes	Yes
Constant	0.612***	0.622***
	(0.133)	(0.136)
Adj. R-squared	0.087	0.085
Kleibergen-Paap Wald F statistic	72.041	27.355
No. of obs.	113,839	113,839

Note: ***, **,* refer to the significance levels at 1, 5 and 10% respectively. Robust standard errors are given in parenthesis. The estimated coefficients for the regional and year dummies are omitted from the table.

2.5.4. 2SLS estimation with parental homeownership as an instrument

Table 2.9 presents the 2SLS estimation results using parental homeownership status as an instrument for the homeownership status with the BCS 1970. Column (1) shows the OLS regression of unemployment status on homeownership status. The next column shows the corresponding IV estimation results. The types of control variables from the BCS 1970 are similar to those from the BHPS but the two datasets are slightly different in terms of types of marital status and housing and number of regions. Unlike the BHPS, there is no information on the number of employed individuals in the household in the BCS 1970.

Parental social class is a new control variable which does not appear in the regression specifications with the BHPS. It needs to be included for the parental homeownership status (the instrument) not to be correlated with their children's unemployment status (the dependent variable). There is the possibility that parents' innate abilities, which would have contributed to their homeownership, are genetically correlated with the abilities of their children. In turn, the children's innate abilities can surely affect their labour market outcomes. Therefore, the parents' homeownership status and children's unemployment status could be spuriously correlated through the parents' innate ability. Hence, the omission of the parental innate abilities in the structural equation (second-stage equation) could correlate the parental homeownership status with their children's unemployment status. As a proxy variable for parents' innate abilities, a father's social class in 1975 when the respondents were 5 years old is included in the regressions. Though there is also the information on a mother's class, it is missing for too many respondents and hence it is better to use the father's class only.

Table 2.9. IV estimation of unemployment with BCS 1970

(Dependent variable: unemployment status (1 if unemployed, otherwise 0))

	(1) OLS	(2) IV
Homeowner (Private renter exc.)	-0.024**	-0.014
	(0.010)	(0.088)
Age	-0.142	0.008
	(0.516)	(0.040)
Age squared	0.002	-0.000
	(0.008)	(0.001)
Male (Female exc.)	0.004	0.004
	(0.003)	(0.003)
Qualification (Higher degree exc.)		
First degree	-0.008	-0.009
	(0.011)	(0.011)
A level	-0.019*	-0.019*
	(0.010)	(0.010)
GCSE	-0.020**	-0.021*
	(0.010)	(0.011)
CSE	-0.019**	-0.020**
	(0.010)	(0.010)
No qualification	-0.016	-0.017
	(0.011)	(0.011)
Marital status (Married exc.)		
Cohabiting	0.005	0.006
	(0.004)	(0.008)
Single (never married)	0.033***	0.034**
	(0.008)	(0.016)
Separated, divorced & widowed	0.024**	0.026
	(0.010)	(0.018)
No. of children	-0.009**	-0.009**
	(0.004)	(0.005)
No. of HH members	0.008**	0.008**
	(0.004)	(0.004)
Father's occupational class when respondent was 5 year old. (Professional occupation exc.)		
Managerial and low professional occupations	0.000	0.001
	(0.005)	(0.005)
Non-manual skilled occupations	0.007	0.007
	(0.006)	(0.006)
Manual skilled occupations	0.007	0.007
	(0.005)	(0.005)
Semi-skilled occupations	0.005	0.005
	(0.006)	(0.006)
Unskilled occupations	0.016	0.016
	(0.013)	(0.013)
Constant	2.535	0.000
	(8.684)	(0.698)
Adj. R-squared	0.022	0.021
No. of obs.	6,118	6,118

Note: ***, **,* refer to the significance levels at 1, 5 and 10% respectively. Robust standard errors are given in parenthesis. The estimated coefficients for the regional dummies are omitted from the table. See Table 2.10 for the related first-stage regression. The endogenous variable is the homeownership status and the instrumental variable is the father's homeownership status when the respondent was 5 year old.

Despite some differences between the two datasets, the OLS estimation produces a very similar result to the one from the BHPS. Homeowners have a lower probability of unemployment by about 2.4% whilst the comparable result from the BHPS regression is 2.6% (column (1) of Table 2.5). Age and age squared do not have any effect on unemployment in this data because there is no variation in age among the cohort members. Some other variables such as the number of dependent children, household size and housing type do not have a substantial effect on the unemployment status. It is likely because the cohort members are in a similar life-cycle stage so that there is not much variation in those variables across the regression sample.

The first-stage regression for the IV estimation in column (2) of Table 2.9 is presented in column (1) of Table 2.10. As expected, the probability of children's homeownership is higher if their parents were homeowners when they were 5 years old and the Kleibergen-Paap Wald F statistic is 18.9. Overall, there is little doubt that parental homeownership status is a good predictor of their children's homeownership status. In column (2) of Table 2.9, the IV estimate of the coefficient of the homeownership status is less negative and statistically not different from zero. This is the same result as the one from the IV estimation with the BHPS. The two different attempts using two different datasets and two different instruments (with different theoretical rationales) still arrive at the same conclusion that there is no statistically significant relationship between one's homeownership and unemployment status.

Table 2.10. First stage regression of IV estimation with BCS 1970
(Dependent variable: homeownership status (1 if homeowner, 0 private renter))

	First stage regression for column (2) of Table 2.9
Age	3.003*
	(1.648)
Age square	-0.045*
	(0.025)
Male (female exc.)	-0.007
	(0.007)
Qualification (Higher degree exc.)	
First degree	0.033*
	(0.020)
A level	0.045**
	(0.018)
GCSE	0.039*
	(0.021)
CSE	0.045**
	(0.019)
No qualification	0.038*
	(0.022)
Marital status (Married exc.)	
Cohabiting	-0.082***
	(0.010)
Single (never married)	-0.149***
	(0.015)
Separated, divorced & widowed	-0.168***
	(0.024)
No. of children	0.014
	(0.009)
No. of HH members	-0.008
	(0.008)
Father's occupational class when respondent was 5 y.o. (Professional occupation exc.)	
Managerial and low professional occupations	-0.008
	(0.014)
Non-manual skilled occupations	0.012
	(0.016)
Manual skilled occupations	0.001
	(0.014)
Semi-skilled occupations	0.005
	(0.016)
Unskilled occupations	0.005
	(0.024)
Father was homeowner at 5 y.o.	0.036***
	(0.008)
Constant	-48.987*
	(27.625)
Adjusted R-squared	0.079
Kleibergen-Paap Wald F statistic	18.890
No. of obs.	6,118

Note: ***, **,* refer to the significance levels at 1, 5 and 10% respectively. Robust standard errors are given in parenthesis. The estimated coefficients for the regional dummies are omitted from the table.

2.6. Conclusion

This chapter attempts to test the causal relationship between homeownership and unemployment. The most well-known argument regarding the relationship is Oswald's hypothesis that homeownership causes unemployment through a reduced mobility effect. This mechanism was initially suggested to explain the persistently high unemployment rate in Europe during the 1960s and onwards and is empirically supported by early macro-level work. However, more recent studies criticise the methodology of macro-studies, focus more on individual outcomes and often find no causal relationship. This chapter contributes to the literature by adding more evidence on the causal relationship using UK micro-data, the BHPS and the BCS 1970.

A major empirical challenge is to remove bias caused by the endogeneity of homeownership which is problematic for two reasons. Firstly, becoming a homeowner is a complicated process in which numerous individual and household characteristics are involved. It is virtually impossible to control for all the relevant factors which affect both homeownership and unemployment status. Secondly, there exists a potential reverse causation from unemployment to homeownership through an income effect. Panel FE models are adopted to deal with a significant part of the omitted variable bias (i.e. the bias introduced by time-invariant unobservables). As a more fundamental solution to the bias caused by the endogenous explanatory variable, an instrumental approach is applied with the instrument being either the local homeownership rate or the parental homeownership status.

The OLS or logit estimations show that homeowners are associated with a lower probability of being unemployed contrary to Oswald's suggestion. It is suspected that this could be because of the endogeneity of homeownership status. As expected, a comparison of a pooled model with a panel FE model indicates that unobserved individual heterogeneity does play a role in determining the homeownership and unemployment statuses simultaneously. When time-constant individual characteristics are controlled for, the difference in the unemployment probability between homeowners and renters disappears. The pooled and panel FE IV estimation also supports the view that homeownership status does not cause unemployment regardless of the instrument employed.

CHAPTER 3. DISTANCE MOVED AND HOUSING TENURE DECISION

3.1. Introduction

This chapter explores whether distance moved affects movers' tenure decisions. The proposition tested, in a nutshell, is that households who move farther away from their original residence are less likely to own their next property since the collection of available information on the destination housing market is more difficult and more costly and hence the mover household may be more likely to make an ill-informed investment decision by either paying too high a price or buying into the 'wrong' neighbourhood. Renters may also misjudge the prevailing rent-level or choose the 'wrong' neighbourhood but this has less grave consequences. One reason for this is that renters who move, in contrast to owners, face much lower transaction costs.

It is well reported in the literature that job-related movers move relatively longer distances than housing-related movers (Nivalainen, 2004). This is because people relocating for work typically move between geographically separate labour markets, which generally also means a new housing market. Therefore, in the new housing market, movers might not have sufficient information on neighbourhood quality, housing stock conditions and housing price level. One interesting question that can be raised regarding the housing market outcomes of inter-city movers is how the difficulty of obtaining information on destination housing markets might affect residential tenure decisions.

In the absence of sufficient information on the destination housing market, movers may prefer private renting over owner-occupation, due to the following distinctive characteristics of housing. Firstly, the collection of information on local housing markets is costly as straightforward comparisons across residential properties and neighbourhoods are not possible due to the heterogeneity of housing units and neighbourhoods in terms of location and characteristics. Secondly, if a house and its neighbourhood in the new housing market turn out to be not as good as the mover initially thought, the costs to bear would be greater for homeowners than for private renters. This is because homeowners invest large proportions of their wealth in housing and thus cannot diversify investment risk effectively (Hilber, 2005). Moreover, their housing investment is heavily leveraged, which magnifies the risk further. Lastly, buying a home involves much higher transaction costs than renting it. If a household fails to purchase desirable housing due to lack of information, they would be either locked into that property, or forced to bear additional transaction costs to relocate.

Given the importance of local housing market information for the home-buyer, the starting point of this chapter is that the amount of information on the destination housing market is likely to decrease with the distance moved. As the distance a mover plans to move increases, searching in the target housing market would become increasingly costly and therefore the amount and quality of information on the housing and neighbourhood is likely to fall.[4] A reasonable strategy for the mover's tenure decision would be to rent a property first and delay a home

[4] The emergence of the Internet may have substantially reduced such costs, but still a site inspection in person is essential for properties, unlike mass produced goods sold online and hence costs of travels for viewing are inevitable.

purchase until she or he gains sufficient knowledge and living experience in the new neighbourhood. If this hypothesis holds true, it should be observed that the longer distance a household moves, the more likely they are to choose private renting over owner-occupation.

For the empirical analysis, this chapter employs the Survey of English Housing (SEH) which provides essential information such as the housing tenure status of households (the dependent variable), distance moved (the key explanatory variable) and various demographic and socio-economic characteristics of household heads (the control variables). One particular concern with the empirical analysis below is that of spurious correlation or omitted variable bias: distance moved may be correlated with propensity to own since omitted characteristics of households may be correlated with both, the dependent and the key explanatory variable. Controlling for as many demographic and socio-economic characteristics of mover households may therefore mitigate potential omitted-variable bias. The additional advantage of the SEH is that it provides some information on the pre-move conditions of households. In particular, the tenure status at the previous accommodation helps to control for unobservable preferences and ability to own of households.

The empirical analysis first looks at whether the awareness of neighbourhood problems decreases with the distance moved. The awareness of neighbourhood problems is assumed to represent how much information mover households had on the new neighbourhood. It turns out that the further the distance moved, the less likely the movers are to be aware of neighbourhood problems, such as crime,

vandalism, litter and graffiti. This result indicates that movers do have difficulty in collecting information on distant housing markets, which leads to the main hypothesis of this chapter regarding whether the probability of choosing owner-occupation over private renting decreases with the distance moved. When a number of demographic and socio-economic factors are controlled for to reduce omitted variable bias, a negative effect of the distance moved on the probability of homeownership is unambiguously demonstrated. Lastly, an investigation is carried out regarding the household's subsequent moving decision after the initial move. Once renting households become familiar with the local housing market over time, they may consider moving to more permanent owner-occupied accommodation. The tendency to move out shortly after the first move is expected to increase with the distance moved. This is because the longer-distance movers are more likely to find problems with the accommodation and neighbourhoods since they were less able to collect information on them prior to the moves. Therefore, the length of stay at the current accommodation is expected to have a negative relationship with the distance moved and this hypothesis is supported by the empirical analysis.

The structure of the chapter is as follows. Section 3.2 discusses previous related studies and the contribution of this chapter to the relevant literature. Section 3.3 derives testable hypotheses using a simple model. Section 3.4 explains the empirical strategies and data. Section 3.5 presents empirical results. The last section concludes with a summary and discussion of policy implications.

3.2. Literature review

A lack of knowledge on a destination housing market essentially refers to uncertainty about its various aspects, such as general housing conditions, neighbourhood quality, trend in housing price moves, etc. Therefore, a review of the literature on whether and how various types of uncertainty affect housing tenure decisions would give us the chance to assess indirectly the effects of informational shortage on tenure decisions. In the literature on determinants of housing tenure, many kinds of uncertainty have received academic attention and have been reported to discourage homeownership. One of the types of uncertainty that has drawn the greatest academic attention is income uncertainty. Since it usually accounts for the largest proportion of mortgage payments of households, uncertainty in the income stream could discourage homeownership. Haurin and Gill (1987) show a negative relationship between a likelihood of homeownership and labour income uncertainty which is measured by the proportion of military personnel household income accounted for by spouse's earnings which are arguably more uncertain in their future prospects. Using the coefficient of variation of income as a measure of income uncertainty, Haurin (1991) shows that a 10% increase in income variability has as negative an effect on homeownership as a 5% decrease in income itself. Robst et al. (1999) obtain a more precise measure of income risk by excluding expected income rises and still reach the same conclusion as the studies above. A similar finding is also reported in the European context where institutional settings and property market characteristics are different from the US. For Germany and Spain, households with a high degree of income uncertainty tend to prefer renting when the uncertainty is represented by both variance and skewness (Diaz-Serrano, 2005).

The level and volatility of housing prices is also found to be important in homeownership decisions. Fu (1991) shows that the causal relationship between housing price uncertainty and probability of home ownership should be negative in the absence of borrowing constraints. However, if borrowing against expected future gains from housing investment is constrained, the relationship may be ambiguous due to the offsetting of income and substitution effects (Fu, 1995). Most empirical studies have reported the negative effect of housing price uncertainty on the probability of becoming a homeowner. Rosen et al. (1984) find that uncertainty involved in the relative prices of owning has significantly lowered the proportion of owner-occupiers in the U.S. Turner (2003) also finds that an increase in anticipated price volatility by one standard deviation lowers the probability of homeownership by 7 percent. Turner and Seo (2007) suggest that when the anticipated volatility in housing price increases from the 25^{th} to the 50^{th} percentile of the distribution of price variance, the probability of transitioning into ownership is reduced by 19 percent.

There are several studies that explore how general knowledge on housing transactions (e.g. how to search housing markets and obtain mortgages from credit markets) affects tenure decisions. Dietz and Haurin (2003) suggest that a positive correlation in the homeownership status between parents and their child could be attributed to a transfer of knowledge on housing transactions. Haurin and Morrow-Jones (2006) argue that some studies (e.g. Henretta, 1984; Boehm and Schlottman, 1999; Mulder and Smits, 1999) have produced empirical outcomes consistent with the intergenerational transfer of housing market knowledge.

Haurin and Morrow-Jones (2006) explore how the market knowledge affects homeownership decision in a more direct manner. Through their own survey, they measure the level of market knowledge of home-buyers based on questions around the market situation at the time of the home-purchase.[5] They find that the degree of real estate market knowledge increases the probability of homeownership and the difference in market knowledge can explain 8.5% of the gap in homeownership rates between black and white households.

The related literature indicates that various types of uncertainty lower the likelihood of homeownership and it can be inferred that a lack of information or knowledge on local housing markets could also discourage homeownership. Since housing is a location-specific good, knowledge on local housing markets may have a greater impact on the homeownership decision than general knowledge of the process of buying a home. The focus of previous studies has been typically on the latter type of knowledge, whereas the focus of this chapter is on the former type of knowledge. The investigation of the relationship between distance moved and housing tenure can have important policy implications for both labour and housing markets. It is therefore surprising that, to the author's knowledge, the topic of this chapter has not been explored *rigorously* so far. The most closely related study to this chapter is Clark and Huang (2004) who look at the relationship between the distance moved and the homeownership status using the UK BHPS. They suggest that homeowners do not show a particularly strong tendency to return to renting even after they make long-distance moves but their

[5] Haurin and Morrow-Jones (2006) collected the following information: (1) whether there was knowledge on mortgage interest rates; (2) the size of the down payment; (3) whether help was available from the government; (4) whether it was possible to find out own, and; (5) other's credit scores, and; (6) whether the affordable housing price was known

conclusion is based solely on descriptive statistics. This chapter is the first study to test the hypothesis formally in a rigorous way through an econometric approach.

3.3. Testable hypotheses

A theoretical model which predicts the main hypothesis of this chapter requires a couple of assumptions. The first assumption is that the amount of information obtained on new housing and neighbourhoods decreases with the distance moved. Though the case is different from this chapter, Yezer and Thurston (1976) and DaVanzo (1983) argue that the longer the distance of the move, the more costly it is to obtain information on employment opportunities in the new labour market. Their finding arguably also applies to the housing market. One may argue that unlike the 1970s and 1980s, the acquisition of information on distant housing markets may not be too difficult any more in an era of highly advanced information technology. However, the types of market information which are electronically available are typically limited to standardised factors, such as housing price, size and structure. There are still many types of crucial information which are difficult to obtain unless properties and neighbourhoods are actually experienced, such as detailed conditions of properties or the neighbourhood atmosphere. Final decisions on whether to buy a property are typically made only after several visits. Therefore, it is predicted that the distance a household wants to move would lower the likelihood of homeownership through its negative effect on efforts to collect information on local housing markets. This assumption is first tested prior to the main hypothesis.

Hypothesis 1

The amount of information held by a household on the new neighbourhood would decrease as the distance moved increases.

The second assumption is that the maximum utility receivable from desirable accommodation and neighbourhood would be greater for homeowners than for private renters. Homeowners are likely to have a higher level of satisfaction from good accommodation as they can stay as long as they want. However, if the accommodation turns out to be of poor quality, owner-occupiers would receive a lower level of utility from it than private renters as the former have to either stay and suffer from the poor housing quality or incur large costs to relocate. In any case, homeowners are likely to suffer more than renters when choosing a wrong property. Mathematically,

$$U_G^H > U_G^R \text{ and } U_B^H < U_B^R \quad (1)$$

Where U_G^H and U_G^R are the utilities that a homeowner and a renter receive from *good* accommodation respectively. U_B^H and U_B^R are the utilities that a homeowner and a renter receive from the *bad* accommodation respectively.

A tenure decision is made by comparing the utilities which a homeowner and a renter are expected to receive from the new accommodation. For simplicity, the expected utility is given as the weighted average utility of the maximum and minimum achievable with weights respectively given by the probability of finding good accommodation (p_d) and bad accommodation (1-p_d). Then, the difference in the expected utility between owner-occupation and renting is given by

$$E(U_{HO}) - E(U_{RT}) = p_d \cdot \left(U_{G,HO} - U_{G,RT}\right) + (1 - p_d) \cdot \left(U_{B,HO} - U_{B,RT}\right) \quad (2)$$

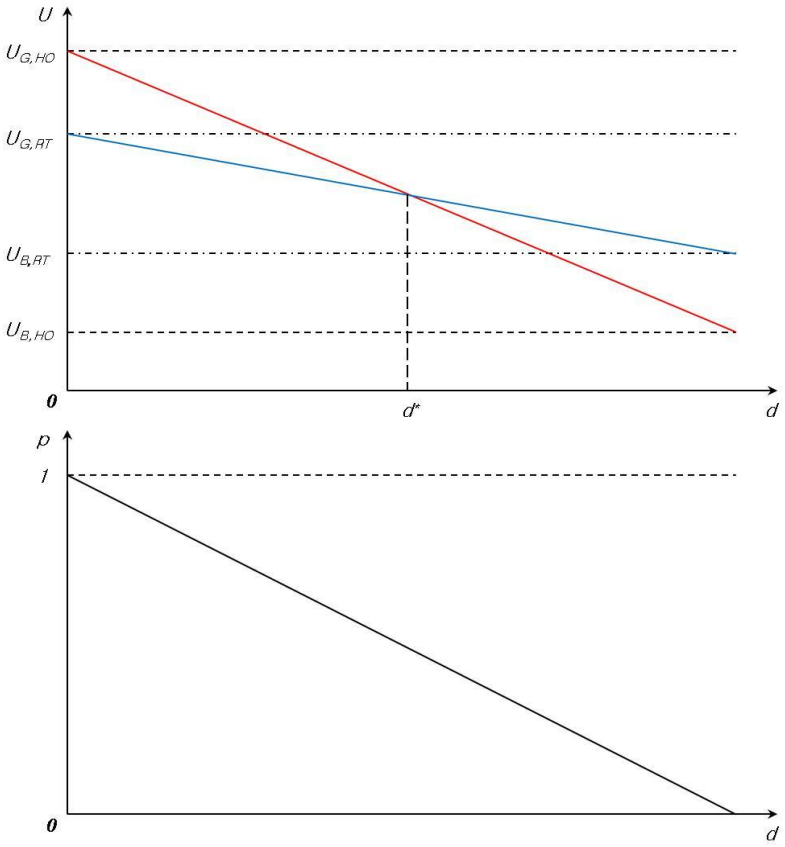
Where $E(U_{HO})$ and $E(U_{RT})$ are the expected utilities that a homeowner and a renter receive from the new accommodation respectively.

If the two assumptions hold true, the difference in the expected utility given in *eq.* (2) is positive when the distance moved is zero, but turns negative eventually as the distance moved increases. For a relatively short distance, the probability (p_d) would be close to 1 since sufficient searching can be done. Hence, the first term on the right-hand side of *eq.* (2) which is positive by assumption is given more weight so that the difference in the expected utility between homeowners and private renters remains positive, implying that owner-occupation is preferred over renting. As the probability (p_d) decreases with the distance moved, a greater weight is given to the second term on the right-hand side of *eq.* (2), which is assumed to be negative, causing the difference $(E(U_{HO}) - E(U_{RT}))$ to be negative and renting is preferred. Therefore, beyond a certain distance from the previous accommodation, where the difference in the expected utility turns negative, movers can be expected to increasingly choose renting over owning.

The argument made above is depicted in Figure 3.1, which contains the assumptions made previously. Since the probability of finding good accommodation (p_d) decreases with the distance moved, both schedules of expected utility for homeowner and renter decrease but only at different rates as the difference in the utility received from good and poor accommodation is greater for homeowners than renters. Therefore, the two schedules should

intersect each other somewhere in the middle between the origin and destination of the move and the intersecting point is referred to by d^*. Between the point of origin and the point d^*, the expected utility is greater for homeowners than it is for renters and hence movers are likely to own, whilst beyond d^*, they are likely to rent.

Figure. 3.1. Expected utility and the intended moving distance.



So, the main hypothesis is derived as follows:

Hypothesis 2

The longer the distance moved, the more likely the households are to choose private renting over owner-occupation.

After settling down in the new accommodation, the level of the mover's knowledge of the local area can be expected to increase over time and then, she or he may want to make a so-called corrective or adjustment move to more permanent accommodation. The tendency to adjust housing quality and location within the same housing market is likely to be stronger for longer-distance movers as they are less likely to be satisfied with their first accommodation in the new market due to informational shortage. The third hypothesis can be expressed as follows:

Hypothesis 3

The longer the distance moved, the more likely movers are to move again shortly after the initial move.

Through these hypotheses, this chapter aims to show that facing uncertainty regarding the quality and conditions of new housing and neighbourhood, a mover's rational behaviour may be to first move into temporary accommodation, such as a privately rented house, and then later settle in a more permanent place.

3.4. Data and empirical strategy

3.4.1. Survey of English Housing

The data for the empirical analysis is the Survey of English Housing (SEH) provided by the Office of National Statistics, the UK governmental body in charge of national statistics. As its title implies, the geographical span of the survey covers England and Wales. The SEH provides all the essential information for the analysis, such as distance moved, housing tenure status, problems in neighbourhoods, housing-related characteristics, demographic and socio-economic characteristics of households. An additional advantage is that it provides some important information on the pre-move conditions of households. Of these, previous tenure status is particularly helpful in controlling for households' preferences and ability to own. Furthermore, the lengthy period of the survey (15 years from 1993/4 to 2007/8) permits securing a large sample size and carrying out experiments for sub-sample groups.[6] The unit of observation is a household and some types of personal information are available mainly for household heads.

3.4.2. Common variables and sample selection

There are some explanatory variables commonly included in Hypothesis 1 to 3. Before empirical strategies for individual hypotheses are considered, it seems worth discussing what those variables are and how they are constructed. More detailed discussion of how each variable is relevant to the hypotheses will be done in the empirical strategy section for each hypothesis. As this chapter seeks to

[6] The SEH ended in 2007/8 and was merged with the English House Condition Survey to form a single housing survey for the UK called the English Housing Survey. Due to the continuity and consistency of the variables, this chapter uses only SEH.

explain the relationship between distance moved and homeownership status, these two variables are discussed first before the other control variables are introduced. A housing tenure is the dependent variable for Hypothesis 2 and one of the explanatory variables for the other hypotheses. In the SEH, housing tenure is given in greater detail, but without sacrificing the efficiency of information they can be grouped into three categories: homeowners, private renters and public renters. The detailed tenure types and their matching groups are shown in Table 3. 1. Of many characteristics of renting, this chapter focuses on its ability to offer easy and quick access to and exit from accommodation. In this regard, the inclusion of public renters in the regression sample appears inappropriate since it is more difficult and takes longer time to secure public housing in the UK due to a high demand for them and the complex administrative process. Rather, it is often chosen as permanent accommodation for low-income households, as an alternative to owner-occupied housing. Therefore, it seems appropriate to exclude public renters from regressions.

Table 3.1. Types of housing tenure

Housing tenure	Detailed categories
Homeowner	own outright, own with mortgage, partly own/partly rent
Public renter	rent from local council or housing association
Private renter	rent from private landlords, property companies, employers, organisation or relatives and friends

The main explanatory variable for Hypothesis 1 to 3 is the distance between current and previous accommodation, as self-assessed by moving households. In the SEH, it is originally given as a categorical variable which has 9 options

ranging from ‘under 1 mile’, ‘1 miles but not 2 miles’, ‘2 miles but not 5 miles’, ‘5 miles but not 10 miles’, ‘10 miles but not 20 miles’ , ‘20 miles but not 50 miles’, ’50 miles or more’, ‘From Northern Ireland’ to ‘From abroad’. The households that moved from Northern Ireland or abroad are excluded from the regression sample as they are likely to choose private renting mainly because they are unfamiliar with the institutional settings of the property and mortgage markets of England and Wales, rather than because they do not have knowledge of the local areas where they have settled down. The inclusion of movers from outside England and Wales would lead to an overestimation of the influence of the distance moved on the tenure decision, unless there are any control variables indicating how knowledgeable these overseas migrants were about England and Wales and the property market in general prior to moving. Rather than using the distance moved as it is given as a categorical variable, it is converted to a continuous variable by taking a mid-value of each range, for example, 0.5 miles for the category of ‘under 1 mile’, 1.5 miles for ‘1 mile but not 2 miles’, 3.5 miles for ‘2 miles but not 5 miles’, etc. For a category of ‘over 50 miles’, ‘75 miles’ is assigned arbitrarily.[7] This way of conversion makes it easy to interpret the empirical results and regressions with interaction terms between distance moved and other characteristics of households are possible.

The rest of the control variables can be grouped into: (1) demographic and individual-specific characteristics of household heads; (2) household-level information; (3) housing characteristics, and; (4) time and regional dummies. The individual-specific characteristics include a household head’s age, sex and

[7] Though not shown in this chapter, other arbitrary numbers such as 50 and 100 miles are assigned for ‘over 50 miles’ category but the overall result does not change accordingly.

economic status. There are 5 types of economic status, as shown in the summary tables below, among which students are not included in the regression sample. As the SEH interviews household heads, the students in the survey are usually those who have left their family for college or university and live independently from their parents. Therefore, they become heads of their own households. In general, they move long distance to the places of study and rent a house. However, their tenure types are likely to be determined mainly by the expected length of stay rather than by the distance moved. Therefore, the inclusion of students in the regression sample would dilute the true relationship between tenure status and distance moved.

The household-level characteristics include the household composition, the number of adults, the number of children and the real household income. The numbers of adults and children indicate the size of the household. The income to be used in the regression is the real annual household income. The nominal household income is first obtained by summing up the annual gross income earned by the household head and spouse. For single-person households, the spouse's income is set to zero. The nominal income is deflated to the real income in 1993 prices, using the UK retail price index. Though the income is initially given as a continuous variable in pound sterling, it enters the regressions as a categorical variable and this allows for regressions using the interaction terms between the distance moved and the income. The categorical household income takes four options of £0-£9,999, £10,000-£19,999, £20,000-£49,999 and £50,000 or over.

The information on housing characteristics covers the type of accommodation, the number of bedrooms, the council tax band and the main reasons for moving. There are five types of accommodation (see the summary tables below for details) and the number of bedrooms is used as a proxy for the size of the accommodation. The council tax band is adopted as a proxy measure for property value. There are 9 categories of the property value ranging from 'up to £40,000' to 'over £320,000' (see the summary tables below for the detailed categories). Though it does not really show the exact value of each property, it reveals the *relative* position of the property in the spectrum of housing value. Another advantage of the band is that it shows the value of rented properties as well as that of owner-occupied ones.

The main reasons for moving are found to be highly correlated with the distance moved and also with the homeownership status. In the SEH, there are more than 20 different reasons for moving and the types and the number of the reasons vary from year to year. To make these consistent over time, each and every individual reason is classified into one of the categories from Table 3.2 below. Those who report 'for homeownership' as the main reason for moving are excluded from the regression sample. This chapter looks at the impact of distance moved on tenure status but since their moves were motivated by the intention to become homeowners, it is unlikely that the distance moved affects homeownership.

The total number of households available for the entire duration of the survey is 429,878. The information on the distance moved is available for those households which moved within 3 years of the time of the survey, meaning they were relatively recent movers, and the number of these cases is 67,648. Out of this

number, public renters, foreign migrants, students and those who moved for homeownership are excluded as discussed above, leaving a sample size of 44,489. When the cases which have a missing value in any of the variables are excluded, the sample size is reduced to 37,755. Starting from this number of observations, the exact sample size used to test each hypothesis varies somewhat due to the exclusion of some cases according to the characteristics of experimentation. This will be explained further when each hypothesis is discussed below.

Table 3.2. Main reasons for moving

Grouped reasons for moving	Stated individual reasons for moving	Year of appearance
Neighbourhood-related reasons	Wanted to move to better and more pleasant neighbourhood	1993-2007
Housing-related reasons	Wanted larger or better housing	1993-2007
	Wanted smaller or cheaper housing	1993-2007
Ownership	Wanted to buy	1993-2007
Involuntary move	Had to leave tied accommodation	1993-1997
	Could not afford mortgage payments	1993-1997
	Could not afford rent payments	1993-1997
	Could not afford mortgage payments or rent payments	1998-2007
	Accommodation was no longer available	1993-1996
	Assured short-hold came to an end	1995-1997
	Landlord required tenant to move out	1995-2007
Personal/family-related reasons	Divorced or separated	1993-2007
	Married or cohabited	1993-2007
	Moved for family reasons	1993
	Moved for personal reasons	1993
	Moved for other family or personal reasons	1994-2007
	Wanted independent accommodation	1994-2007
Job-related reasons	Wanted to move near to new jobs	1993-1997
	Wanted to move near to current job	1993-1997
	Moved for job related reasons	1998-2007
Other reasons	Went to or finished college/university	1993-1997
	Moved for better schools for children	2005-2007
	Other reasons	1993-2007

3.4.3. Empirical strategy

3.4.3.1. Hypothesis 1

Hypothesis 1 is concerned with the relationship between the distance moved and the level of information held by the movers on the new neighbourhood. The SEH provides information on how serious household heads think their local area problems such as crime are. The seriousness of crime is likely to be felt only after people become familiar with their neighbourhood through a few years' residency. Therefore, how aware they are of the local problems is considered closely related to the level of knowledge on their local areas. Therefore, using the degree of awareness of local problems as a proxy for local knowledge, the model specification is written as follows:

$$\textit{Awareness of local problem} = f(\textit{distance moved, other controls}) \qquad (3)$$

The dependent variable takes one of the three options – a specific local problem, for example, crime is 'serious', 'problematic but not serious' or 'not a problem at all'. The ordered logit model is appropriate for the estimation of *eq*. (3), as only an *order* among the options of the dependent variables is known. If this hypothesis is true, longer-distance movers are expected to be less aware of their neighbourhood problems. To check whether the hypothesis holds true for different types of problems, this chapter looks at four types of local problems, namely, crime, vandalism, litter and graffiti.

The seriousness of the local problems felt by movers is affected not only by the distance moved, but also by the quality of neighbourhoods. If neighbourhood

quality is correlated with the distance moved but omitted from the regressions, the estimated coefficient of the distance moved would be biased. In reality, the correlation between the distance moved and neighbourhood quality can occur for two reasons. Firstly, households may want to move away from local problems through long distance moves. Secondly, good-quality neighbourhoods are relatively rare and hence one may have to move relatively long distances to find them. Unfortunately, neighbourhood quality is difficult to measure and there is no information on neighbourhood quality in the SEH. Therefore, the best strategy available is to control for individual household characteristics that capture the quality of neighbourhoods. Two obvious characteristics are real household income and housing value. Real household income is highly associated with neighbourhood quality due to the fact that the higher the income, the more likely the household is to live in the good neighbourhood. Housing value also has a close relationship with neighbourhood quality through the process of housing price capitalisation. Properties in high-quality neighbourhoods are more expensive than those in low-quality ones, reflecting the difference in the neighbourhood quality. Also, the size of housing may be related to the quality of neighbourhood. Large (and expensive) houses are found more often in high-quality neighbourhoods. The number of bedrooms will be controlled for as a proxy for the size of houses.

Hypothesis 1 is concerned with the extent to which mover households were aware of neighbourhood problems *at the time of moves*. As the awareness of the problems tends to go up with the length of residency in the neighbourhood, the regression sample used to test Hypothesis 1 should be limited to mover

households that had lived for relatively a short period in their current accommodation when they were surveyed. However, if the sample is limited to households with too short length of residency, say, one month, it will be difficult to secure a sufficient sample size to obtain reliable empirical results. Therefore, it seems reasonable to choose only those households which had lived in their current houses for less than 1 year (henceforth, first-year residents). Out of 37,775 cases (see 'sample selection' section above for how this number was reached), 15,202 had lived in their locations for less than 1 year. In addition, the sample size varies depending on the types of neighbourhood problems, as, for example, is the sample size is 10,543 for crime. The summary statistics for the variables included in the regressions for Hypothesis 1 are shown in Table 3.3 below.

Table 3.3. Summary statistics for Hypothesis 1

Variable	Obs.	Mean	Std. Dev.	Min	Max
Crime	10543	2.495	0.657	1	3
Distance moved	10543	16.317	25.205	0.5	75
Age	10543	36.671	12.928	16	91
Age squared	10543	1511.908	1172.708	256	8281
Sex	10543	1.294	0.455	1	2
Economic status (Full-time employed exc.)					
Part-time employed	10543	0.070	0.255	0	1
Unemployed	10543	0.032	0.176	0	1
Retired	10543	0.061	0.239	0	1
Inactive	10543	0.066	0.248	0	1
Household composition (Single exc.)					
Couple	10543	0.601	0.490	0	1
Lone parent	10543	0.088	0.283	0	1
Multi-family HH	10543	0.057	0.231	0	1
No. of children	10543	0.553	0.920	0	7
No. of adults	10543	1.800	0.705	1	8
Real HH income (£0-£9,999 exc.)					
£10,000-£19,999	10543	0.293	0.455	0	1
£20,000-£49,000	10543	0.401	0.490	0	1
£50,000 or more	10543	0.074	0.261	0	1
Homeowner (private renter exc.)	10543	0.597	0.491	0	1
No. of bedrooms	10543	2.579	1.004	1	9
Accommodation type (Detached/Bungalow exc.)					
Semi-detached	10543	0.260	0.439	0	1
Terraced	10543	0.309	0.462	0	1
Purpose-built flat	10543	0.125	0.331	0	1
Converted flat	10543	0.113	0.317	0	1
Council tax band (Up to 40k exc.)					
Up to £52k	10543	0.206	0.404	0	1
Up to £68k	10543	0.220	0.414	0	1
Up to £88k	10543	0.174	0.379	0	1
Up to £120k	10543	0.094	0.292	0	1
Up to £160k	10543	0.048	0.213	0	1
Up to £320k	10543	0.036	0.185	0	1
Over £320k	10543	0.007	0.081	0	1

Table 3.3. Summary statistics for Hypothesis 1 (cont.)

Variable	Obs.	Mean	Std. Dev.	Min	Max
Main reasons for moving (Neighbourhood-related exc.)					
Housing-related	10543	0.265	0.441	0	1
Had to move	10543	0.135	0.342	0	1
Personal/family-related	10543	0.349	0.477	0	1
Job-related	10543	0.130	0.336	0	1
Region (North East exc.)					
North West	10543	0.117	0.321	0	1
Yorkshire & the Humber	10543	0.098	0.297	0	1
East Midlands	10543	0.087	0.282	0	1
West Midlands	10543	0.084	0.278	0	1
Eastern	10543	0.117	0.322	0	1
London	10543	0.137	0.343	0	1
South East	10543	0.192	0.394	0	1
South West	10543	0.120	0.325	0	1
Year of survey (1994 exc.)					
1995	10543	0.042	0.202	0	1
1997	10543	0.110	0.313	0	1
1998	10543	0.121	0.326	0	1
1999	10543	0.110	0.313	0	1
2001	10543	0.111	0.314	0	1
2002	10543	0.105	0.306	0	1
2003	10543	0.098	0.297	0	1
2004	10543	0.096	0.294	0	1
2005	10543	0.083	0.276	0	1
2006	10543	0.087	0.282	0	1

3.4.3.2. Hypothesis 2

The main goal of this chapter is to test whether the distance moved affects the mover's tenure decision. The regression model is specified such that the probability of homeownership is expressed as a function of the distance moved along with other control variables.

$$Pr\ (homeownership=1) = f\ (distance\ moved,\ other\ controls) \quad (4)$$

The dependent variable is a dummy which takes the value of 1 if a household is an owner-occupier or 0 if it rents from a private landlord. The model is estimated by maximum likelihood logit. If the hypothesis is true, the estimated coefficient of the distance moved should take a negative sign. The control variable which appears uniquely for Hypothesis 2 is tenure status in the previous accommodation, which takes one of the three options - homeowner, public renter and private renter. This past tenure status is likely correlated with both the current tenure status and the distance moved. Previous homeowners tend to become homeowners again and are reluctant to move short distances as substantial moving costs cannot be justified for short-distance moves.

In order to test Hypothesis 2, only first-year residents are included in the regression sample. This chapter is interested in the relationship between the distance moved and the homeownership status *at the time of the moves* but it will be diluted over time as long-distance movers, who are more likely to be renters, are expected to move again at a *faster* rate than short-distance movers. Therefore, the regression sample needs to be confined to relatively recent mover households.

As counted previously, the first-year residents are 15,202. When the cases with missing values relating to previous tenure are excluded, the final sample size for the regressions is 13,185. See Table 3.4 for the relevant summary statistics.

Table 3.4. Summary statistics for Hypothesis 2

Variable	Obs.	Mean	Std. Dev.	Min	Max
Homeownership status	13185	0.607	0.489	0	1
Distance moved	13185	17.174	25.929	0.5	75
Age	13185	38.515	13.123	17	92
Age squared	13185	1655.622	1221.500	289	8464
Sex	13185	1.271	0.445	1	2
Economic status (Full-time employed exc.)					
Part-time employed	13185	0.068	0.252	0	1
Unemployed	13185	0.038	0.190	0	1
Retired	13185	0.073	0.260	0	1
Inactive	13185	0.072	0.259	0	1
Household composition (Single exc.)					
Couple	13185	0.619	0.486	0	1
Lone parent	13185	0.092	0.289	0	1
Multi-family HH	13185	0.053	0.223	0	1
No. of children	13185	0.626	0.971	0	7
No. of adults	13185	1.830	0.712	1	8
Real HH income (£0-£9,999 exc.)					
£10,000-£19,999	13185	0.284	0.451	0	1
£20,000-£49,000	13185	0.401	0.490	0	1
£50,000 or more	13185	0.079	0.269	0	1
Previous tenure type (Previous HO exc.)					
Previous public renter	13185	0.065	0.246	0	1
Previous private renter	13185	0.406	0.491	0	1
No. of bedrooms	13185	2.653	1.018	1	10
Accommodation type (Detached/Bungalow exc.)					
Semi-detached	13185	0.266	0.442	0	1
Terraced	13185	0.289	0.453	0	1
Purpose-built flat	13185	0.115	0.318	0	1
Converted flat	13185	0.112	0.315	0	1

Table 3.4. Summary statistics for Hypothesis 2 (cont.)

Variable	Obs.	Mean	Std. Dev.	Min	Max
Council tax band (Up to 40k exc.)					
Up to £52k	13185	0.195	0.396	0	1
Up to £68k	13185	0.220	0.414	0	1
Up to £88k	13185	0.186	0.389	0	1
Up to £120k	13185	0.106	0.308	0	1
Up to £160k	13185	0.053	0.225	0	1
Up to £320k	13185	0.041	0.199	0	1
Over £320k	13185	0.008	0.089	0	1
Main reasons for moving (Neighbourhood-related exc.)					
Housing-related	13185	0.312	0.463	0	1
Had to move	13185	0.136	0.343	0	1
Personal/family-related	13185	0.280	0.449	0	1
Job-related	13185	0.134	0.340	0	1
Region (North East exc.)					
North West	13185	0.115	0.319	0	1
Yorkshire & the Humber	13185	0.089	0.285	0	1
East Midlands	13185	0.084	0.277	0	1
West Midlands	13185	0.083	0.276	0	1
Eastern	13185	0.120	0.325	0	1
London	13185	0.141	0.348	0	1
South East	13185	0.202	0.401	0	1
South West	13185	0.123	0.328	0	1
Year of survey (1993 exc.)					
1994	13185	0.062	0.242	0	1
1995	13185	0.069	0.253	0	1
1996	13185	0.068	0.252	0	1
1997	13185	0.075	0.263	0	1
1998	13185	0.083	0.276	0	1
1999	13185	0.075	0.264	0	1
2000	13185	0.075	0.264	0	1
2001	13185	0.074	0.262	0	1
2002	13185	0.071	0.256	0	1
2003	13185	0.067	0.250	0	1
2004	13185	0.065	0.246	0	1
2005	13185	0.057	0.232	0	1
2006	13185	0.059	0.236	0	1
2007	13185	0.058	0.234	0	1

3.4.3.3. Hypothesis 3

The third hypothesis states that the longer the distance moved, the more likely movers are to make an adjustment move shortly after the initial move. If it holds true, the proportion of the second- and third-year residents would be lower among the longer-distance movers than among the shorter-distance ones. Then, the model specification can be written as follows.

$$\textit{Length of stay} = f(\textit{distance moved, other controls}) \quad (5)$$

The dependent variable is the length of stay in the current accommodation. It is given as a categorical variable taking one of the three options: ‘less than 1 year’, ‘1 year but not 2 years’ and ‘2 years but not 3 years’. The reason for why the length of stay (the dependent variable) does not go beyond 3 years is because the information on the distance moved (the key explanatory variable) is only available for those who have lived for less than 3 years in the current accommodation. The categorical dependent variable is converted into a continuous variable by taking the mid-value of the range (e.g. a half year for ‘less than 1 year’). Since it is now a continuous variable, the OLS estimation is applied to *eq*. (5) and the estimated coefficient of the distance moved is expected to have a negative sign if the hypothesis is correct.

In general, the set of other control variables is similar to those used for the previous hypotheses. Of the explanatory variables, the tenure types are particularly important, as private renters are more likely to move again soon and also to have moved longer distances than homeowners. Therefore, unless the types

of housing tenure are controlled for, the estimate for the distance moved would be biased downwards. Also, the main reasons for moving would affect both the distance moved and the length of stay greatly. For example, housing- or neighbourhood-related movers are expected to move short distances and stay relatively long in the same place, while job-related movers would move longer distances and stay only short-term.

Unlike the previous hypotheses, Hypothesis 3 does not require the regression sample to be limited to the first-year residents, as the length of stay is the main explanatory variable. The entire 37,755 cases (see ‘sample selection’ part above) are used and Table 3.5 below presents the relevant summary statistics.

Table 3.5. Summary statistics for Hypothesis 3

Variable	Obs.	Mean	Std. Dev.	Min	Max
Length of stay	37755	1.878	0.818	1	3
Distance moved	37755	16.003	24.873	0.5	75
Age	37755	38.829	13.206	16	95
Age squared	37755	1682.124	1241.046	256	9025
Sex	37755	1.259	0.438	1	2
Economic status (Full-time employed exc.)					
Part-time employed	37755	0.062	0.242	0	1
Unemployed	37755	0.029	0.168	0	1
Retired	37755	0.075	0.263	0	1
Inactive	37755	0.060	0.237	0	1
Household composition (Single exc.)					
Couple	37755	0.653	0.476	0	1
Lone parent	37755	0.081	0.273	0	1
Multi-family HH	37755	0.040	0.195	0	1
No. of children	37755	0.647	0.972	0	7
No. of adults	37755	1.842	0.685	1	9
Real HH income (£0-£9,999 exc.)					
£10,000-£19,999	37755	0.289	0.453	0	1
£20,000-£49,000	37755	0.422	0.494	0	1
£50,000 or more	37755	0.081	0.273	0	1
Homeowner (private renter exc.)	37755	0.715	0.451	0	1
No. of bedrooms	37755	2.702	0.989	1	10
Accommodation type (Detached/Bungalow exc.)					
Semi-detached	37755	0.284	0.451	0	1
Terraced	37755	0.291	0.454	0	1
Purpose-built flat	37755	0.104	0.305	0	1
Converted flat	37755	0.086	0.281	0	1
Council tax band (Up to 40k exc.)					
Up to £52k	37755	0.196	0.397	0	1
Up to £68k	37755	0.219	0.414	0	1
Up to £88k	37755	0.195	0.396	0	1
Up to £120k	37755	0.107	0.309	0	1
Up to £160k	37755	0.054	0.226	0	1
Up to £320k	37755	0.040	0.195	0	1
Over £320k	37755	0.008	0.090	0	1

Table 3.5. Summary statistics for Hypothesis 3 (cont.)

Variable	Obs.	Mean	Std. Dev.	Min	Max
Main reasons for moving (Neighbourhood-related exc.)					
Housing-related	37755	0.295	0.456	0	1
Had to move	37755	0.125	0.331	0	1
Personal/family-related	37755	0.328	0.469	0	1
Job-related	37755	0.120	0.325	0	1
Region (North East exc.)					
North West	37755	0.123	0.328	0	1
Yorkshire & the Humber	37755	0.095	0.294	0	1
East Midlands	37755	0.084	0.277	0	1
West Midlands	37755	0.088	0.283	0	1
Eastern	37755	0.120	0.325	0	1
London	37755	0.129	0.335	0	1
South East	37755	0.193	0.395	0	1
South West	37755	0.120	0.324	0	1
Year of survey (1993 exc.)					
1994	37755	0.050	0.218	0	1
1995	37755	0.059	0.235	0	1
1996	37755	0.060	0.238	0	1
1997	37755	0.072	0.259	0	1
1998	37755	0.084	0.278	0	1
1999	37755	0.078	0.268	0	1
2000	37755	0.080	0.271	0	1
2001	37755	0.078	0.268	0	1
2002	37755	0.075	0.263	0	1
2003	37755	0.073	0.261	0	1
2004	37755	0.069	0.253	0	1
2005	37755	0.066	0.248	0	1
2006	37755	0.061	0.240	0	1
2007	37755	0.057	0.232	0	1

3.5. Empirical results

3.5.1. Local information and distance moved

Table 3.6 presents the ordered logit regression results for Hypothesis 1 that local housing market information decreases with the distance moved (i.e., the underlying proposed mechanism that is driving long-distance movers to rent rather than own). The dependent variable is a measure of how serious household heads think crime is in their local area. Since the dependent variable takes a lower value out of 1, 2 and 3 if household heads think more strongly that crime is serious, the estimate of the distance coefficient would have a positive sign if Hypothesis 1 holds true. The explanatory variables are controlled for in a gradual manner from left to right. In the first column, when the distance moved and regional and time dummies are in the regression model, the estimate turns out to be positive and significant at the 1% level, which is consistent with the hypothesis. However, it is likely to be a biased estimate due to omitted variables.

When household heads' personal and household characteristics are controlled for in column (2), the estimate for the distance moved goes down (and arguably becomes less biased). This is not surprising since various demographic and socio-economic characteristics are correlated with the distance moved to access better neighbourhood quality. For example, retired households move about 19 miles on average when they relocated to access better neighbourhood quality, whereas the average distance moved to access better neighbourhood quality is only 10 miles. Though not explicitly shown in Table 3.6, the addition of real household income contributes further to the decline of the estimated coefficient on 'distance moved', consistent with the prediction made previously.

Table 3.6. Ordered logit estimation of neighbourhood problem awareness 1 (Dependent variable: Seriousness of crime in the area)

	(1) Distance moved only	(2) Personal/HH char.	(3) Tenure, housing char.	(4) Main reasons for moving
Distance moved	0.00530***	0.00449***	0.00406***	0.00343***
	(0.00083)	(0.00084)	(0.00085)	(0.00102)
Age		0.003	-0.007	-0.011
		(0.011)	(0.011)	(0.011)
Age squared		0.000	0.000	0.000*
		(0.000)	(0.000)	(0.000)
Sex (Female exc.)		0.054	0.052	0.051
		(0.053)	(0.054)	(0.054)
Economic status of HH (Full-time employed exc.)				
Part-time employed		-0.116	-0.137	-0.138
		(0.088)	(0.089)	(0.090)
Unemployed		-0.319**	-0.288**	-0.285**
		(0.127)	(0.127)	(0.127)
Retired		0.070	-0.002	-0.012
		(0.156)	(0.157)	(0.158)
Inactive		-0.315***	-0.288***	-0.291***
		(0.099)	(0.099)	(0.100)
Household composition (Single exc.)				
Couple		0.285***	0.210***	0.190**
		(0.076)	(0.078)	(0.078)
Lone parent		0.107	0.023	0.032
		(0.095)	(0.096)	(0.097)
Multi-family HH		0.348***	0.271**	0.242*
		(0.125)	(0.126)	(0.126)
Number of children		-0.028	-0.051*	-0.057**
		(0.025)	(0.027)	(0.027)
Number of adults		-0.173***	-0.188***	-0.189***
		(0.045)	(0.047)	(0.047)
Household real income (£0-£9,999 exc.)				
£10,000-£19,999		0.078	0.013	0.017
		(0.068)	(0.069)	(0.069)
£20,000-£49,999		0.230***	0.092	0.094
		(0.073)	(0.076)	(0.076)
£50,000 or more		0.149	0.005	0.011
		(0.098)	(0.108)	(0.108)
Homeowner (Private renter exc.)			-0.110**	-0.109**
			(0.050)	(0.051)
Number of bedrooms			-0.052*	-0.052*
			(0.031)	(0.031)
Accommodation type (Detached/bungalow exc.)				
Semi-detached			-0.066	-0.069
			(0.067)	(0.067)

Table 3.6. Ordered logit estimation of neighbourhood problem awareness 1 (cont.)

	(1) Distance moved only	(2) Personal/HH char.	(3) Tenure, housing char.	(4) Main reasons for moving
Terraced			-0.313***	-0.297***
			(0.072)	(0.072)
Purpose-built flat			-0.281***	-0.260***
			(0.096)	(0.096)
Converted flat			-0.650***	-0.639***
			(0.102)	(0.103)
Council tax band (Up to £40k exc.)				
Up to £52k			0.295***	0.295***
			(0.066)	(0.066)
Up to £68k			0.392***	0.387***
			(0.069)	(0.069)
Up to £88k			0.392***	0.393***
			(0.081)	(0.081)
Up to £120k			0.352***	0.346***
			(0.095)	(0.096)
Up to £160k			0.419***	0.409***
			(0.121)	(0.122)
Up to £320k			0.450***	0.437***
			(0.138)	(0.138)
Over £320k			-0.059	-0.061
			(0.225)	(0.225)
Main reasons for moving (Neighbourhood-related exc.)				
Housing-related				-0.325***
				(0.071)
Had to leave				-0.357***
				(0.081)
Personal/family-related				-0.382***
				(0.070)
Job-related				-0.244***
				(0.092)
Constant	0.714***	0.966***	0.315	-0.084
	(0.129)	(0.258)	(0.282)	(0.292)
Government office regions	Yes	Yes	Yes	Yes
Years of survey	Yes	Yes	Yes	Yes
Pseudo R-squared	0.0217	0.0302	0.0375	0.0392
No. of obs.	10543	10543	10543	10543

Note: ***, **,* significance level at 1, 5, 10% respectively. Age, sex and economic status are of household heads.

Column (3) additionally controls for housing characteristics. Adding the various housing controls reduces the coefficient on 'distance moved' further. Though not shown explicitly in Table 3.6, a gradual control of housing characteristics reveals that most housing controls do not significantly alter the 'distance moved'

coefficient. The exception is the housing value control (derived from council tax bands), which reduces the coefficient on ‘distance moved’ substantially. Lastly, when the main reasons for moving are also controlled for, the estimate of interest decreases further, as shown in column (4). It implies that the reasons for moving are correlated with both the neighbourhood quality and the distance moved. For example, movers who relocate for neighbourhood-related reasons are likely to move relatively short distances while movers who relocate for job-related reasons usually move long distances and are likely to care less about neighbourhood quality. Importantly, however, even after carefully controlling for a large set of covariates in column (4), the estimated coefficient of the distance moved still remains positive and highly statistically significant at the 1% level.

Table 3.7 shows the results for the same regression as that reported in column (4) of Table 3.6, but with alternative dependent variables: awareness of vandalism, graffiti and litter. Though the results vary somewhat depending on the outcome measure, the estimates of the distance moved variable are positive and statistically significant throughout. All the results are consistent with the hypothesis that the distance moved does have a negative effect on the level of information on destination housing markets held by the movers.

Table 3.7. Ordered logit estimation of neighbourhood problem awareness 2 (Dependent variables: Seriousness of vandalism, graffiti and litter)

Dependent variable	(1) Vandalism	(2) Graffiti	(3) Litter
Distance moved	0.00382***	0.00251*	0.00203**
	(0.00106)	(0.00135)	(0.00100)
Age	0.003	-0.007	-0.025**
	(0.011)	(0.014)	(0.011)
Age squared	-0.000	0.000	0.000**
	(0.000)	(0.000)	(0.000)
Sex (Female exc.)	0.054	0.077	-0.054
	(0.054)	(0.066)	(0.052)
Economic status of HH (Full-time employed exc.)			
Part-time employed	-0.172**	-0.174	-0.246***
	(0.087)	(0.107)	(0.087)
Unemployed	-0.041	-0.148	-0.169
	(0.122)	(0.149)	(0.122)
Retired	0.268*	-0.084	-0.316**
	(0.160)	(0.218)	(0.155)
Inactive	-0.239**	-0.059	-0.258***
	(0.097)	(0.123)	(0.099)
Household composition (Single exc.)			
Couple	-0.005	0.031	0.005
	(0.080)	(0.099)	(0.078)
Lone parent	0.088	0.134	0.015
	(0.097)	(0.123)	(0.097)
Multi-family HH	-0.060	0.034	-0.130
	(0.127)	(0.155)	(0.125)
Number of children	-0.064**	-0.048	-0.025
	(0.028)	(0.034)	(0.027)
Number of adults	-0.070	-0.012	-0.073
	(0.048)	(0.060)	(0.047)
Household real income (£0-£9,999 exc.)			
£10,000-£19,999	0.053	0.189**	-0.062
	(0.068)	(0.085)	(0.067)
£20,000-£49,999	0.192**	0.185**	-0.026
	(0.076)	(0.094)	(0.075)
£50,000 or more	0.293**	0.070	0.038
	(0.116)	(0.135)	(0.114)
Homeowner (Private renter exc.)	-0.039	-0.190***	-0.057
	(0.051)	(0.065)	(0.050)
Number of bedrooms	-0.039	-0.124***	-0.114***
	(0.032)	(0.041)	(0.032)
Accommodation type (Detached/bungalow exc.)			
Semi-detached	-0.171**	-0.088	-0.111
	(0.074)	(0.098)	(0.075)
Terraced	-0.458***	-0.404***	-0.650***
	(0.077)	(0.101)	(0.077)

Table 3.7. Ordered logit estimation of neighbourhood problem awareness 2 (cont.)

Dependent variable	(1) Vandalism	(2) Graffiti	(3) Litter
Purpose-built flat	-0.418***	-0.398***	-0.433***
	(0.102)	(0.130)	(0.101)
Converted flat	-0.671***	-0.630***	-0.846***
	(0.107)	(0.133)	(0.105)
Council tax band (Up to £40k exc.)			
Up to £52k	0.326***	0.335***	0.401***
	(0.065)	(0.080)	(0.063)
Up to £68k	0.523***	0.503***	0.610***
	(0.071)	(0.088)	(0.070)
Up to £88k	0.496***	0.524***	0.679***
	(0.082)	(0.104)	(0.081)
Up to £120k	0.630***	0.691***	0.791***
	(0.101)	(0.128)	(0.102)
Up to £160k	0.597***	0.731***	1.099***
	(0.129)	(0.164)	(0.136)
Up to £320k	0.741***	0.966***	1.310***
	(0.152)	(0.187)	(0.159)
Over £320k	0.912***	1.668***	0.855***
	(0.295)	(0.426)	(0.286)
Main reasons for moving (Neighbourhood-related exc.)			
Housing-related	-0.267***	-0.338***	-0.357***
	(0.077)	(0.099)	(0.076)
Had to leave	-0.309***	-0.341***	-0.398***
	(0.086)	(0.110)	(0.084)
Personal/family-related	-0.330***	-0.355***	-0.358***
	(0.076)	(0.097)	(0.074)
Job-related	-0.326***	-0.298**	-0.350***
	(0.098)	(0.125)	(0.095)
Constant	-0.935***	-1.712***	-1.955***
	(0.290)	(0.360)	(0.283)
Government office regions	Yes	Yes	Yes
Years of survey	Yes	Yes	Yes
Pseudo R-squared	0.0332	0.0440	0.0482
No. of obs.	11594	10360	11689

Note: ***, **,* significance level at 1, 5, 10% respectively. Age, sex and economic status are of household heads.

3.5.2. Homeownership status and distance moved

Table 3.8 shows the results of logit estimates for Hypothesis 2 (the main proposition), which states that the longer the distance moved, the lower the probability of being a homeowner. The explanatory variables that are thought to be correlated with both the distance moved and the probability of homeownership are grouped into three categories and controlled for gradually, from left to right. In

all columns, the distance moved has a negative relationship with the probability of homeownership, which is consistent with Hypothesis 2. However, the estimates for the distance moved vary substantially depending on the types of control variables included in the regressions. When personal, household and housing characteristics are controlled for, the negative relationship between distance moved and the probability of homeownership becomes stronger, as shown in columns (2) and (3).

However, a dramatic fall in the estimate of distance moved occurs when the main reasons for moving are controlled for in column (4). This implies that those reasons are correlated with the distance moved and homeownership status. More specifically, they are correlated with the distance moved in a way that those who want to move for job-related reasons (often for better job opportunities in large labour markets such as London) need to move long distances whilst those moving for housing- or neighbourhood-related reasons move short distances. This is thought to be because new housing and neighbourhoods are likely to be available within a shorter distance from previous accommodation than large job markets are. This is confirmed by simple statistics as shown in Table 3.9. The average distance moved for housing-related movers is 5 to 6 miles, for neighbourhood-related 13 to 17 miles and for job-related about 50 miles among the sample used for the regressions in Table 3.8.

Table 3.8. Logit estimation of homeownership status (Dependent variable: Homeownership status)

	(1) Distance moved only	(2) Personal/HH char.	(3) Housing char.	(4) Main reasons for moving
Distance moved	-0.00527***	-0.0108***	-0.0140***	-0.00538***
	(0.00069)	(0.00084)	(0.00095)	(0.00118)
Age		0.124***	0.013	0.010
		(0.012)	(0.013)	(0.013)
Age squared		-0.001***	0.000	0.000
		(0.000)	(0.000)	(0.000)
Sex (Female exc.)		0.365***	0.331***	0.300***
		(0.060)	(0.069)	(0.069)
Economic status of HH (Full-time employed exc.)				
Part-time employed		-0.274***	-0.388***	-0.448***
		(0.091)	(0.101)	(0.102)
Unemployed		-1.116***	-0.851***	-1.033***
		(0.129)	(0.134)	(0.138)
Retired		0.940***	0.709***	0.518***
		(0.177)	(0.179)	(0.184)
Inactive		-1.195***	-1.046***	-1.135***
		(0.100)	(0.105)	(0.107)
Household composition (Single exc.)				
Couple		0.832***	0.728***	0.687***
		(0.084)	(0.095)	(0.097)
Lone parent		0.002	-0.392***	-0.369***
		(0.097)	(0.106)	(0.107)
Multi-family HH		-0.536***	-0.685***	-0.712***
		(0.139)	(0.160)	(0.162)
Number of children		0.156***	-0.148***	-0.147***
		(0.026)	(0.030)	(0.030)
Number of adults		-0.139***	-0.355***	-0.365***
		(0.049)	(0.053)	(0.054)
Household real income (£0-£9,999 exc.)				
£10,000-£19,999		0.861***	0.669***	0.703***
		(0.071)	(0.079)	(0.080)
£20,000-£49,999		1.468***	0.937***	0.987***
		(0.076)	(0.088)	(0.090)
£50,000 or more		1.748***	0.832***	0.832***
		(0.109)	(0.136)	(0.137)
Previous tenure status (Previous HO exc.)				
Previous public renter			-1.184***	-1.285***
			(0.092)	(0.093)
Previous private renter			-1.815***	-2.019***
			(0.054)	(0.061)

Table 3.8. Logit estimation of homeownership status(cont.)

	(1) Distance moved only	(2) Personal/HH char.	(3) Housing char.	(4) Main reasons for moving
Number of bedrooms			0.374***	0.370***
			(0.039)	(0.039)
Accommodation type (Detached/bungalow exc.)				
Semi-detached			0.010	-0.002
			(0.082)	(0.084)
Terraced			-0.190**	-0.210**
			(0.087)	(0.090)
Purpose-built flat			-0.940***	-0.912***
			(0.113)	(0.115)
Converted flat			-1.418***	-1.410***
			(0.120)	(0.123)
Council tax band (Up to £40k exc.)				
Up to £52k			0.266***	0.254***
			(0.078)	(0.078)
Up to £68k			0.467***	0.456***
			(0.083)	(0.084)
Up to £88k			0.572***	0.561***
			(0.095)	(0.097)
Up to £120k			0.359***	0.422***
			(0.115)	(0.117)
Up to £160k			0.603***	0.617***
			(0.149)	(0.152)
Up to £320k			0.329*	0.382**
			(0.178)	(0.184)
Over 320k			0.171	0.302
			(0.356)	(0.357)
Main reasons for moving (Neighbourhood-related exc.)				
Housing-related				-0.094
				(0.077)
Had to leave				0.632***
				(0.095)
Personal/family-related				-0.300***
				(0.081)
Job-related				-1.130***
				(0.105)
Constant	0.881***	-3.646***	-0.040	0.330
	(0.127)	(0.294)	(0.356)	(0.368)
Government office regions	Yes	Yes	Yes	Yes
Years of survey	Yes	Yes	Yes	Yes
Pseudo R-squared	0.0134	0.2131	0.3581	0.3771
No. of obs.	13185	13185	13185	13185

Note: ***, **,* significance level at 1, 5, 10% respectively. Age, sex and economic status are of household heads.

Table 3.9. Average distance moved by residential tenure status

Reasons for moving	Average distance moved (miles)		Proportion by tenure (%)	
	Home-owner	Private renter	Home-owner	Private renter
Total sample	16.03	18.94	60.67	39.33
Those who moved				
for neighbourhood	17.08	13.60	70.64	29.36
for housing	6.73	5.10	71.51	28.49
forced to move	8.40	7.34	53.59	46.41
for personal/family affair	20.03	16.17	56.17	43.83
for jobs	50.11	51.31	41.74	58.26

Note: The sample used to compute the figures in this table are the same as those used in the regressions of which results are presented in Table 3.8.

At the same time, the reasons for moving are also correlated with the homeownership status. The housing- and neighbourhood-related movers tend to owner-occupy their new homes but the job-related movers tend to rent, *regardless of the distance moved*. Compared to housing- and neighbourhood-related movers, job-related movers would be less certain about where exactly they would settle down in the long term within the new housing market as they might not have collected sufficient information prior to the move as their primary concerns were jobs and job locations. On the other hand, housing- and neighbourhood-related movers will have obtained sufficient information on the new housing and the neighbourhood, so that they can make better informed tenure decisions and are therefore more prone to homeownership, when other things, including the distance moved, are equal. In fact, this is another piece of evidence that the information available on the local housing market has an influence on movers' housing tenure decisions. Overall, the reasons for moving are correlated with both distance moved and tenure decisions and controlling for the reasons substantially reduces

the strength of the negative correlation between them. However, the correlation is still negative and significant at the 1% level, consistent with Hypothesis 2.

Another, potentially important, variable determining the housing tenure status is financial constraints or the ability to buy a home. Financially constrained households may not be able to owner-occupy regardless of the distance moved. In a setting where all households are financially constrained, nobody would own and the distance moved would be unrelated to homeownership. Therefore, the more financially constrained households are, the more biased towards zero the relationship between the distance moved and homeownership can be expected to be. Household income, one of the controls used in the analysis, partly captures the ability of households to afford homeownership, as it reflects the degree to which households are liquidity constrained. The SEH, like most other household datasets, does not include, however, household wealth, which is a measure for how down payment constrained households are. Whereas the SEH does not include a direct measure of how wealthy households are, the dataset does include information on the previous housing tenure status, that is, whether households have collateral (i.e., a home), to proceeds of which they can use to purchase a home in a subsequent move. Previous homeowners are also better placed to obtain a new mortgage. In other words, previous homeowners are less likely to be financially constrained when moving and considering whether to buy or rent at the new place. The interaction effects between different types of previous tenure (homeowner, private renter and public renter), reported in column (1) of Table 3.10, reveal that distance moved indeed only affects the tenure choice at the destination location of previous homeowners but not of previous private renters or social renters, arguably because

the latter two categories are financially more constrained. Another implication from this result is that informational constraints matter even for those who are rich enough to buy housing. Even if households want to and are able to owner-occupy, homeownership is likely to be deterred if they are uncertain about the quality of potential housing and the surrounding environment. They may choose to first rent and then buy at the new destination location. The results are indicative that lack of housing and neighbourhood information are very important for housing tenure decisions.

Table 3.10. Logit estimation of homeownership status with interaction terms (Dependent variable: Homeownership status)

	(1) Distance × Prev. tenure	(2) Distance × HH income	(3) Distance × Reasons for moving
Interaction term:			
Distance moved × Previous tenure			
Distance×prev. HO	-0.00957***		
	(0.00141)		
Distance×prev. public renter	-0.00377		
	(0.00341)		
Distance ×prev. private renter	0.00014		
	(0.00164)		
Interaction term:			
Distance moved × HH income			
Distance×£0-9,999		0.00045	
		(0.00201)	
Distance×£10,000-19,999		-0.00289	
		(0.00191)	
Distance×£20,000-49,999		-0.00996***	
		(0.00168)	
Distance×£50,000 or more		-0.0129***	
		(0.00328)	
Interaction term:			
Distance moved × Reasons for moving			
Distance×neighbourhood			0.00123
			(0.00306)
Distance×housing			0.00192
			(0.00371)
Distance×had to move			-0.01080***
			(0.00386)
Distance×personal/family reasons			-0.00368**
			(0.00187)
Distance×job-related reasons			-0.01147***
			(0.00223)

Table 3.10. Logit estimation of homeownership status with interaction terms (cont.)

	(1) Distance × Prev. tenure	(2) Distance × HH income	(3) Distance × Reasons for moving
Age	0.009	0.010	0.008
	(0.013)	(0.013)	(0.013)
Age squared	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Sex (Female exc.)	0.295***	0.297***	0.295***
	(0.069)	(0.069)	(0.069)
Economic status of HH (Full-time employed exc.)			
Part-time employed	-0.440***	-0.440***	-0.455***
	(0.103)	(0.103)	(0.102)
Unemployed	-1.018***	-1.044***	-1.052***
	(0.139)	(0.139)	(0.138)
Retired	0.544***	0.489***	0.477**
	(0.184)	(0.183)	(0.185)
Inactive	-1.117***	-1.144***	-1.145***
	(0.107)	(0.107)	(0.107)
Household composition (Single exc.)			
Couple	0.697***	0.676***	0.671***
	(0.097)	(0.097)	(0.097)
Lone parent	-0.378***	-0.348***	-0.370***
	(0.108)	(0.107)	(0.107)
Multi-family HH	-0.683***	-0.717***	-0.736***
	(0.162)	(0.162)	(0.162)
Number of children	-0.148***	-0.145***	-0.143***
	(0.030)	(0.030)	(0.030)
Number of adults	-0.369***	-0.359***	-0.360***
	(0.054)	(0.054)	(0.054)
Household real income (£0-£9,999 exc.)			
£10,000-£19,999	0.700***	0.741***	0.708***
	(0.080)	(0.089)	(0.080)
£20,000-£49,999	0.994***	1.154***	1.000***
	(0.090)	(0.098)	(0.090)
£50,000 or more	0.844***	1.101***	0.851***
	(0.136)	(0.165)	(0.137)
Previous tenure status (Previous HO exc.)			
Previous public renter	-1.411***	-1.294***	-1.283***
	(0.109)	(0.093)	(0.093)
Previous private renter	-2.216***	-2.032***	-2.031***
	(0.073)	(0.061)	(0.061)
Number of bedrooms	0.371***	0.368***	0.371***
	(0.039)	(0.039)	(0.040)
Accommodation type (Detached/bungalow exc.)			
Semi-detached	-0.007	-0.005	0.010
	(0.084)	(0.085)	(0.085)

Table 3.10. Logit estimation of homeownership status with interaction terms (cont.)

	(1) Distance × Prev. tenure	(2) Distance × HH income	(3) Distance × Reasons for moving
Terraced	-0.206**	-0.211**	-0.202**
	(0.089)	(0.090)	(0.090)
Purpose-built flat	-0.903***	-0.905***	-0.902***
	(0.115)	(0.115)	(0.115)
Converted flat	-1.398***	-1.409***	-1.406***
	(0.124)	(0.124)	(0.124)
Council tax band (Up to £40k exc.)			
Up to £52k	0.250***	0.248***	0.257***
	(0.079)	(0.078)	(0.079)
Up to £68k	0.451***	0.441***	0.454***
	(0.084)	(0.084)	(0.084)
Up to £88k	0.554***	0.554***	0.557***
	(0.097)	(0.097)	(0.098)
Up to £120k	0.414***	0.415***	0.423***
	(0.117)	(0.117)	(0.117)
Up to £160k	0.622***	0.610***	0.612***
	(0.151)	(0.152)	(0.152)
Up to £320k	0.383**	0.373**	0.380**
	(0.182)	(0.184)	(0.184)
Over 320k	0.322	0.265	0.271
	(0.359)	(0.355)	(0.353)
Main reasons for moving (Neighbourhood-related exc.)			
Housing-related	-0.105	-0.092	-0.036
	(0.078)	(0.078)	(0.087)
Had to leave	0.677***	0.640***	0.786***
	(0.096)	(0.096)	(0.107)
Personal/family-related	-0.306***	-0.306***	-0.223**
	(0.081)	(0.081)	(0.095)
Job-related	-1.141***	-1.030***	-0.718***
	(0.103)	(0.106)	(0.150)
Constant	0.447	0.278	0.284
	(0.368)	(0.369)	(0.370)
Government office regions	Yes	Yes	Yes
Years of survey	Yes	Yes	Yes
Pseudo R-squared	0.3785	0.3785	0.3782
No. of obs.	13185	13185	13185

Note: ***, **,* significance level at 1, 5, 10% respectively. Age, sex and economic status are of household heads.

Next, it is examined whether the relationship of interest varies among different income groups. It is predicted that the higher the household income, the less likely they are to be financially constrained and therefore the higher the household income, the more negative the estimated coefficient on distance moved. There are

four levels of household income: ‘£0 to £9,999’; ‘£10,000 to £19,999’; ‘£20,000 to £49,999’, and; ‘£50,000 and over’. Using the interaction terms between the income and the distance moved, the coefficient of the distance moved for each income group is estimated. The empirical results are consistent with the prediction, as the lowest income group has the smallest estimate (with a positive sign), which is not statistically different from zero, while the highest group has the largest and the most statistically significant negative estimate (see column (2) of Table 3.10). The result confirms that only when there is little financial constraint, households can express their preference over residential tenure according to the distance moved.

As discussed previously, the main reasons for moving are correlated with the intensity of the search for housing markets. Those who want to move for better housing or neighbourhood can be expected to search more intensively than those moving for job-related reasons, regardless of the distances they plan to move. Therefore, the relationship between the probability of homeownership and the distance moved can be expected to be weaker and less significant for housing- and neighbourhood-related movers. This hypothesis is tested using the interaction terms between the distance moved and the reasons for moving (column (3) of Table 3.10). The estimates for ‘distance×neighbourhod’ and ‘distance×housing’ are positive but not different from zero statistically, while the rest are negative and significant. Of these, the estimate for the job-related movers is the most negative, and compared to the rest of the movers, job-related movers’ tenure decisions are more strongly affected by the distance moved.

Table 3.11 reports quantitative effects of the relationship of interest based on the various regression results from Table 3.8 and Table 3.10. The first row shows the relationship for the entire sample, indicating the predicted probability of homeownership (67.27%), when the hypothetical distance moved is 0 miles. It goes down with the distance moved and those who moved 75 miles have a 9.41% point lower chance of being a homeowner (see the last column). This suggests that the impact of distance moved on homeownership is quantitatively meaningful. When this is examined by distinguishing between previous tenure categories (Panel A), previous homeowners show the sharpest fall in the probability of becoming a homeowner when their distances moved increase from 0 miles to 75 miles and the difference is statistically significant at the 1% level. The difference in propensity to own as a consequence of moving 75 miles further away is 11.63% point, again quantitatively quite meaningful. When the effect is investigated by real household income (Panel B), the change in the probability of homeownership along the distance moved is the largest for the households with earnings of '£50,000 or more' (-17.33% point). Finally, in Panel C, job-related movers are the group showing the most substantial fall in the likelihood of homeownership (-20.58% point).

Table 3.11. Predicted probability of homeownership by distance moved (%)

	Distance moved				Δ% between 0 and 75 miles
	0 mile	25 miles	50 mils	75 miles	
Total sample	67.27	64.25	61.10	57.86	9.41
Panel A: By previous tenure					
Previous homeowner	84.87	81.54	77.66	73.24	11.63
Previous public renter	33.76	31.69	29.69	27.76	6.00
Previous private renter	33.71	33.78	33.86	33.94	-0.23
Panel B: By household income					
£0-£9,999	46.85	47.12	47.40	47.68	-0.83
£10,000-£19,999	76.49	75.16	73.79	72.37	4.12
£20,000-£49,999	80.16	75.91	71.07	65.69	14.47
£50,000 or more	84.17	79.37	73.58	66.84	17.33
Panel C: By reason for moving					
Neighbourhood-related	66.56	67.24	67.91	68.58	-2.02
Housing-related	64.56	65.65	66.73	67.79	-3.23
Had to leave	78.90	74.06	68.55	62.46	16.44
Personal/family-related	61.95	59.75	57.52	55.26	6.69
Job-related	52.11	44.96	38.02	31.53	20.58

3.5.3. Corrective move and distance moved

Table 3.12 shows the OLS regression results from the test of Hypothesis 3, which states that the longer distance a household has moved, the more likely it is to move again soon to find more permanent accommodation within the same housing market. As discussed previously, the empirical strategy is to run OLS regressions with the dependent variable being length of stay in the current location and the main explanatory variable being distance moved. The estimates of the distance moved would take a negative sign.

Table 3.12. OLS estimation of length of stay
(Dependent variable: length of stay)

	(1) Distance moved only	(2) Personal/HH char.	(3) Housing char.	(4) Distance × Tenure
Distance moved	-0.00093***	-0.00129***	-0.00114***	
	(0.00017)	(0.00017)	(0.00020)	
Interaction terms:				
Distance moved				
×Tenure type				
Distance×Homeowner				-0.00086***
				(0.00024)
Distance×Private renter				-0.00172***
				(0.00029)
Age		0.030***	0.023***	0.023***
		(0.002)	(0.002)	(0.002)
Age squared		-0.000***	-0.000***	-0.000***
		(0.000)	(0.000)	(0.000)
Sex (Female exc.)		0.014	-0.003	-0.003
		(0.012)	(0.011)	(0.011)
Economic status of HH				
(Full-time employed exc.)				
Part-time employed		-0.030	-0.008	-0.009
		(0.019)	(0.018)	(0.018)
Unemployed		-0.067**	0.018	0.015
		(0.026)	(0.026)	(0.026)
Retired		0.030	-0.006	-0.008
		(0.029)	(0.029)	(0.029)
Inactive		-0.078***	0.009	0.007
		(0.021)	(0.020)	(0.020)
Household composition				
(Single exc.)				
Couple		0.059***	0.019	0.018
		(0.016)	(0.016)	(0.016)
Lone parent		0.002	0.013	0.012
		(0.020)	(0.020)	(0.020)
Multi-family HH		-0.069***	-0.018	-0.020
		(0.027)	(0.027)	(0.027)
Number of children		0.048***	0.043***	0.043***
		(0.005)	(0.005)	(0.005)
Number of adults		-0.015*	-0.006	-0.005
		(0.009)	(0.009)	(0.009)
Household real income				
(£0-£9,999 exc.)				
£10,000-£19,999		0.059***	-0.003	-0.002
		(0.014)	(0.014)	(0.014)
£20,000-£49,999		0.113***	0.017	0.019
		(0.015)	(0.016)	(0.016)
£50,000 or more		0.126***	0.043*	0.044*
		(0.021)	(0.022)	(0.022)
Homeowner (Private renter exc.)			0.343***	0.328***
			(0.011)	(0.012)
Number of bedrooms			-0.001	-0.001
			(0.006)	(0.006)

Table 3.12. OLS regressions for Length of stay (cont.)

	(1) Distance moved only	(2) Personal/HH char.	(3) Housing char.	(4) Distance × Tenure
Accommodation type (Detached/bungalow exc.)				
Semi-detached			-0.015	-0.014
			(0.013)	(0.013)
Terraced			-0.023	-0.022
			(0.015)	(0.015)
Purpose-built flat			-0.041**	-0.040**
			(0.020)	(0.020)
Converted flat			-0.056***	-0.056***
			(0.021)	(0.021)
Council tax band (Up to £40k exc.)				
Up to £52k			0.016	0.017
			(0.014)	(0.014)
Up to £68k			0.000	0.001
			(0.015)	(0.015)
Up to £88k			0.016	0.016
			(0.017)	(0.017)
Up to £120k			-0.016	-0.015
			(0.020)	(0.020)
Up to £160k			-0.057**	-0.057**
			(0.025)	(0.025)
Up to £320k			-0.085***	-0.085***
			(0.029)	(0.029)
Over 320k			-0.089*	-0.088*
			(0.051)	(0.051)
Main reasons for moving (Neighbourhood-related exc.)				
Housing-related			-0.015	-0.014
			(0.014)	(0.014)
Had to leave			0.006	0.006
			(0.017)	(0.017)
Personal/family-related			0.012	0.012
			(0.014)	(0.014)
Job-related			0.054***	0.057***
			(0.018)	(0.018)
Government office regions	Yes	Yes	Yes	Yes
Years of survey	Yes	Yes	Yes	Yes
Constant	1.991***	1.093***	1.090***	1.100***
	(0.029)	(0.054)	(0.060)	(0.060)
Adj. R-squared	0.0041	0.0383	0.0675	0.0677
No. of obs.	37755	37755	37755	37755

Note: ***, **,* significance level at 1, 5, 10% respectively. Age, sex and economic status are of household heads.

Indeed, when only the distance moved is controlled for besides the time and regional dummies, its estimate is negative and statistically different from zero at the 1% level. This means that longer-distance movers move again at a faster rate than their shorter-distance counterparts. This result is persistent regardless of the type of control variable included in the regressions. Moreover, the effect of interest is likely to be stronger among the private renters because those who planned to move to more permanent accommodation have become a tenant first. Therefore, the estimate of the distance should be more negative for tenants. This hypothesis is supported by the empirical analysis which examines the effect by type of tenure using the interaction terms between the distance moved and the tenure types, as shown in column (4) of Table 3.12. Though the effect is smaller, even homeowners tend to stay for a shorter period, the longer distances they have moved. This latter finding is indicative of long distance movers generally being more mobile. However, the fact that the effect is about twice as strong for private renters, provides support for the main proposition put forward in this chapter that long distance moves deter homeownership for information related reasons; subsequent to the relocation, long distance movers appear to gather information about the local housing market, which enables them subsequently, to attain homeownership.

3.6. Conclusion

This chapter investigates the relationship between the probability of homeownership and the distance moved. As the distance moved increases, the search for new accommodation becomes increasingly costly in terms of time and money. Without sufficient information on properties in the destination housing market, a household is likely to avoid homeownership due to high housing investment risks and transactions costs. As a result, the probability of homeownership is expected to decrease with the distance moved. It is only after movers become more familiar with the local property markets that they consider making 'corrective' moves to more permanent owner-occupied accommodation. Therefore, it is predicted that the longer the distance moved, the higher the likelihood of moving soon.

Empirically, this chapter tests a series of hypotheses consistent with the prediction above. Firstly, it tests whether a negative relationship exists between the seriousness of the neighbourhood problems felt by households and the distance moved. The degree of the seriousness is seen as how aware household heads are of neighbourhood problems and hence is likely related to the level of local housing market knowledge held by them. The findings in this chapter suggest that, after controlling for factors capturing 'objective' neighbourhood quality, the household-reported seriousness of area problems decreases with the distance moved. This implies that collection of information on the new neighbourhood becomes more difficult and costly as the distance movers plan to move increases.

The main hypothesis of this chapter is that the longer the distance moved, the lower the probability of becoming a homeowner. Those who need to make a long-distance move are likely to avoid homeownership in the new neighbourhood. This is thought to be because long-distance movers cannot collect enough information on the new housing market, based on which they would make tenure decisions. This main proposition finds strong support in the data. The strongest adverse effect of distance moved on homeownership can be found among the least constrained households financially, consistent with theory. When the effect of interest is examined by previous tenure status and by household income, it is the most statistically significant for previous homeowners and the highest income group respectively. An investigation of the effect of interest by reasons for moving reveals that it is not significantly different from zero for housing- and neighbourhood-related movers, as they are likely to pay good attention to housing and neighbourhood regardless of the distance they move, so perhaps incur higher information costs prior to moving. Finally, the hypothesis of 'corrective' moves is tested, which states that the longer the distance moved, the sooner households move again to more permanent accommodation within the same housing market. Empirically, the relationship between the length of stay and the distance moved turns out to be negative and is stronger for private renters, consistent with theory.

Through testing a series of hypotheses, this chapter reveals the importance and relevance to individuals' homeownership decisions of local housing market information as represented by the distance moved. The difficulty of collecting information on destination housing markets deters homeownership. Typically, long-distance moves are found to be associated with job opportunities. When they

move for jobs, they would not normally pay much attention to the nature of accommodation. Even though they were given job opportunities, they may have to let them go if they cannot quickly find places to stay in the new labour markets. Some may use a strategy of ‘moving first and searching chances’ and they still require temporary accommodation, until it is clear where they will work and live. Therefore, the availability of temporary and flexible accommodation such as private renting lessens the worry and effort to find places to stay for job-related movers.

CHAPTER 4. COMMUTING TIME AND WORKER'S EFFORT

4.1. Introduction

This chapter explores whether the length of commuting time can affect a worker's effort in the workplace. For most workers, commuting is an unavoidable daily activity to connect housing and labour markets and accounts for a relatively large proportion of working hours. The RAC foundation (2007) reports that the average commuting time per day is 54 minutes in the UK which would be equivalent to one-eighth of the working hours of full-time workers. It is much higher in large cities such as London where workers spend 84 minutes every working day. Unfortunately, workers' feelings about this important daily activity are generally negative. Rouwendal and Meijer (2001) argue that workers dislike commuting as the value they put on commuting time is even higher than their wage rates. Transportation studies as well as psychological literature also demonstrate the negative effect of commuting on both the physical and mental health of commuters (Koslowskyet et al., 1995; Evans and Wener, 2006; Hoehner et al., 2012). Considering the negative effects of commuting on workers, it is natural to link commuting time with worker effort. Commuting may not only cause commuters to feel exhausted but also impose mental stress from aggravating events such as traffic jams, road accidents, etc. Indeed, it has been reported that the degree of commuting stress negatively impacts on the level of job strain (Wener et al., 2006).

The topic of this chapter, however, has been largely overlooked until recently by economists. Zenou (2002) assumes that workers' effort levels decrease with

commuting distance in order to propose the 'red-lining' hypothesis to explain the supposed spatial mismatch problem first raised by Kain (1968) but does not show whether the assumption is valid empirically. Ha (2005) investigates the same issue using UK micro-data but does not find a statistically significant relationship between effort and commuting time. Van Ommeren and Gutiérrez-i-Puigarnau (2011) study the relationship between commuting distance and worker productivity. Considering that worker productivity is an outcome of underlying effort, the nature of their study very much reflects the current chapter. Using German data, they confirm that commuting distance does increase worker absenteeism rates, as average absenteeism would be lower by 15 to 20% if commuting distances were negligible. This chapter explores the relationship using UK data employing a strategy of careful control with explanatory variables. While Van Ommeren and Gutiérrez-i-Puigarnau (2011) use one proxy measure of productivity (absenteeism), this chapter adopts unpaid overtime work as an additional measure of effort to make sure the empirical results are robust.

Though it is rather an abstract concept, effort is regarded as *effective* labour supply and closely related with labour productivity in the literature. In essence, the two concepts are more or less the same and interchangeable as workers are truly productive only during their effective working hours. As worker effort can be seen as labour supply, this chapter uses a neoclassical labour supply model to make predictions regarding the effect of commuting on the supply of effort by workers. The model has an advantage in that it can accommodate effort in the form of effective *working hours*. The neoclassical model predicts that commuting time negatively affects worker effort: when contracted hours of working are fixed

and a higher wage cannot be paid to longer-distance commuters, the workers would have an incentive to shirk.

Through the review of previous studies in the labour-supply literature, absenteeism and unpaid overtime working hours are chosen as proxy measures for work effort. Absenteeism is assumed to be a negative proxy measure for underlying effort and hence is expected to have a positive relationship with commuting time. However, in practice, absenteeism may have a couple of shortcomings as a proxy variable for effort. Firstly, absenteeism is likely unwelcomed and suppressed by employers and hence it may underestimate the true effort level. Secondly, absenteeism is often caused by random and unexpected illness and injury which are largely unrelated to a worker's effort. Therefore, an additional measure of effort needs to be adopted in the hope that the use of two different measures can strengthen the reliability of the empirical evidence if they lead to the same result. Unpaid overtime hours are another popular measure of effort found in the literature and expected to be negatively correlated with commuting time. Both of these are thought to be valid measures as they bear a core characteristic of effort that it is costly to employees but beneficial to employers.

The data to be used for the empirical analysis is the UK Quarterly Labour Force Survey (QLFS) which provides information on both absenteeism and unpaid overtime work and other relevant variables. The model specifications express absenteeism and unpaid overtime as a function of commuting time and other control variables respectively and are estimated by OLS. The overall empirical

evidence is supportive of the hypothesis. The absenteeism rate turns out to have a positive correlation with commuting time. Unpaid overtime work also has the expected correlation with commuting time: it decreases as commuting time increases.

For the robustness of these results, the sizes of the estimated coefficient of commuting time are compared between sub-groups of the regression sample. Firstly, it is predicted that the relationship between commuting time and effort level would be stronger for part-time than full-time workers. For a given increase in commuting time, the effective wage rate would fall more strongly for part-time workers as their hours of work are shorter. Hence, the decrease in their effort levels would be expected to be more pronounced. Consistent with the prediction, part-time workers lower their effort levels at a faster rate for a given increase in commuting time. Secondly, working women are predicted to reduce effort to a greater extent than working men as commuting time increases. They usually do more household work and hence they are more time-constrained than men. This tendency would be expected to increase were they to have dependent children. As a result, they would respond more sensitively to a given increase in commuting time so that they would reduce their work effort to a greater extent. Indeed, this hypothesis turns out to be supported by the empirical analysis.

Thus, effort level and commuting time are shown to have a negative relationship, as predicted by the theoretical model. However, it is still problematic to assert that the relationship is causal. This chapter suggests that commuting time decreases worker effort but the reverse causation is theoretically possible, as less work-

oriented workers may choose long commutes to live in suburban areas. Therefore, the identification of causality of interest cannot be claimed until more reliable methods to deal with reverse causality, such as an instrumental variable approach, are applied. However, it still seems feasible to propose which direction of causality contributes more to the correlation found in this chapter. Theoretically, less work-oriented workers may want to avoid places of high job density, such as the central business districts, but in reality, they are likely to try to find jobs near their homes since they are less work-motivated. Therefore, the influence of causality from work effort to commuting time is likely limited. So, even if there might be a gap between the true causality of interest and correlation found in this chapter, it would arguably not be too large to invalidate the conclusions of this chapter.

The structure of this chapter hereafter is as follows. Section 4.2 presents a neoclassical labour supply model to predict the effect of commuting time on effort. Section 4.3 reviews selected papers in the literature to identify good proxy measures for effort and its determinants identified empirically. Then, it describes the data and key dependent variables and discusses the empirical strategy. Section 4.4 reports the empirical results and discusses their implications. Section 4.5 concludes with a summary and policy implications.

4.2. Theoretical framework and predictions

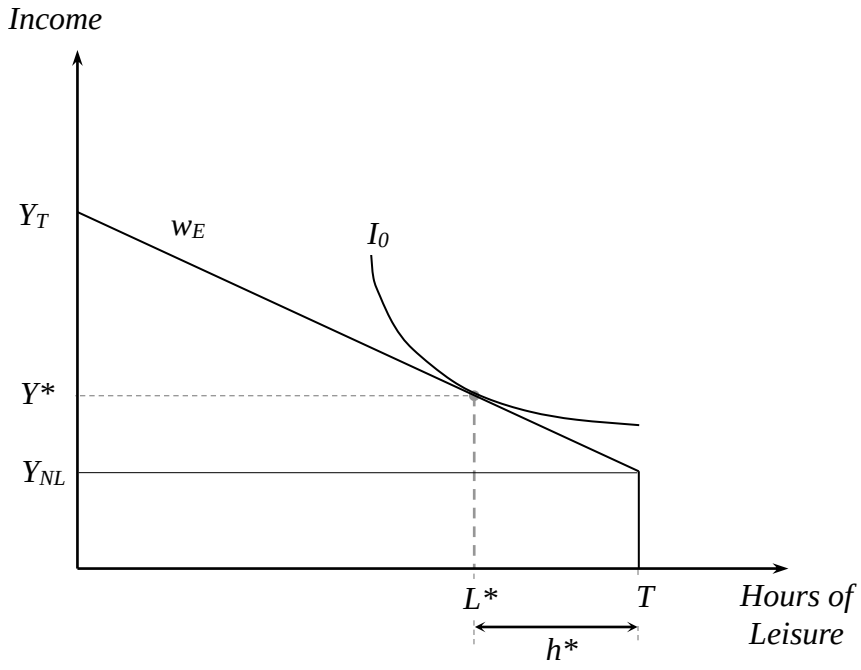
Though the term 'effort' is widely used in labour economics, it is difficult to define and measure due to its abstract and intangible nature. Instead, a popular view on effort is that it would be more or less the same as labour productivity. The well-known efficiency wage or shirking model states that the efficiency wage is paid by firms "to motivate workers and boost their productivity" (Rocheteau, 2000, p. 76). Zenou (2002) also regards effort as labour productivity since his model is grounded in the efficiency wage model. Strictly speaking, however, effort is unlikely to be identical to labour productivity, but rather one of the most influential determinants of labour productivity. Even capital as a production factor helps to increase labour productivity, but they are not the same.

A more conceptually correct view is that effort is an *effective* labour supply as defined by Filer et al. (1996). In reality, the presence of workers in the workplace does not automatically mean that labour has been truly supplied, as what actually contributes to production is the effort which the workers make towards their jobs. For example, it is well-known that the trade union practice of 'working to rule' tends to greatly reduce labour productivity. The extent to which workers engage in shirking behaviour would drive a wedge between contracted labour supply and actual effort exuded. The recognition of effort as true labour supply has an advantage in that it can easily fit in to the existing analytical framework: the neoclassical labour supply model.

Though it is simplistic, the neoclassical labour supply model is useful in capturing the main features of individuals' labour supply decisions and its predictions are

matched by stylised facts in the labour market. The presentation of the neoclassical model follows the one in Borjas (2005) and it would be modified to accommodate commuting time. As depicted in Figure 4.1, the neoclassical model consists of a worker's utility function and budget constraint and shows how the worker determines his preferred working hours by maximising the utility subject to the budget constraint.

Figure 4.1. Labour supply decision of non-commuter

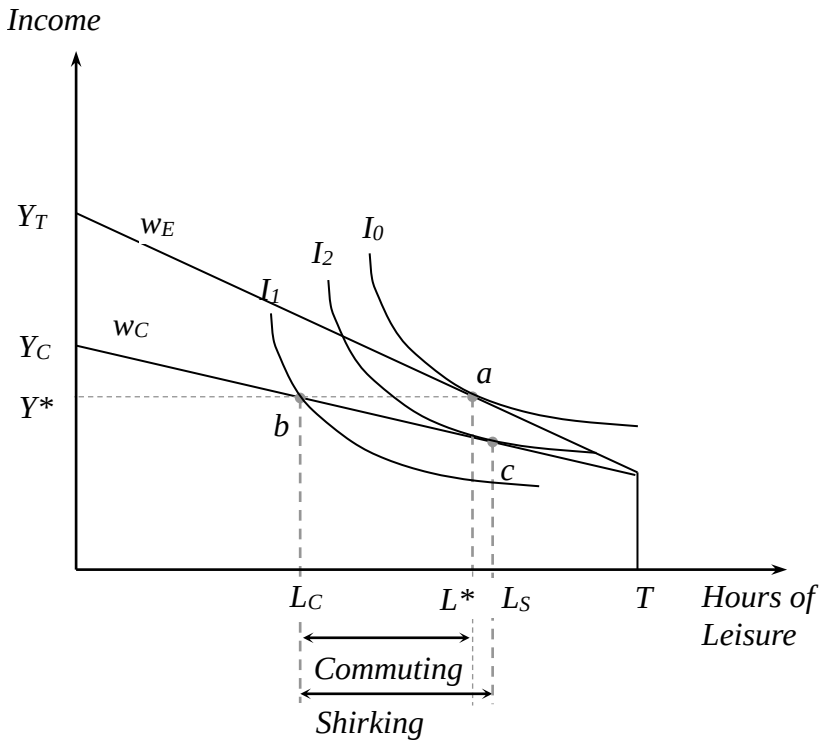


The model assumptions are as follows. A representative worker derives his utility from income (Y) and leisure time (L) which is the remainder of total hours (T) after work. The worker's preference follows the neoclassical axioms of completeness, transitivity, continuity and convexity. As a result, the utility function can be represented by a family of indifference curves which are

continuous, downward-sloping and convex to the origin, which is shown by the indifference curve (I_0) in Figure 4.1. The budget constraint is given by the worker's total income, which can be divided vertically between non-labour income (Y_{NL}) and labour income (i.e. hourly wage rate times hours of work). The negative slope of the budget constraint reflects a trade-off between leisure hours and income and represents a market-determined efficiency wage (w_E), for which a worker is expected to put in maximum effort, as suggested by Shapiro and Stiglitz (1984). Optimal working hours can be determined at the tangential point between the indifference curve and budget constraint. At the point of tangency, the hours of leisure are L^*, corresponding income is Y^*, and hours of work are h^*. As the worker is paid the efficiency wage, he is working for h^* with full effort.

Unlike the original neoclassical model where commuting activity is not considered, the modified version explicitly recognises commuting time in Figure 4.2. Commuting is a grey area in the sense that it is neither leisure time nor working time. Although leisure refers to any activities other than the supply of labour, it should be confined to those which generate positive utility to the worker. In this sense, commuting cannot be a part of leisure as it is generally believed to cause 'bads' to workers and hence reduce utility. Commuting is rather an action closely tied to work and thus workers may consider the time spent on commuting as a part of implicit working hours.

Figure 4.2. Labour supply decision of commuter



If commuting time is included in working hours, the effective working time becomes the hours of work plus commuting time and the recalculated leisure time is now represented by L_C in Figure 4.2. From the point of view of employers, however, commuting is the worker's own responsibility and hence they have no reason to pay for it. As a result, there is no change in the worker's labour income and it stays the same at Y^*. Therefore, with positive commuting time ($L^* - L_C$), the worker has to be at point *b* and the new budget constraint should pass through this point. The slope of the new budget constraint, which is associated with the wage rate (w_C), is obviously less steep than the original one (w_E) because the same income (Y^*) needs to be spread over extended working hours ($T - L_C$), which are longer than the original working hours ($T - L^*$) determined in Figure 4.1.

At point *b*, the utility level is represented by the indifference curve, I_1 and lower than the level represented by I_0 at point *a*. This implies that positive commuting time results in a contraction in the opportunity set and decline in the utility level as expected. Given that the wage rate is now w_C, point *b* is not a place where the worker's utility is maximised. The worker's optimal behaviour would be to decrease the working time by ($L_S - L_C$) and stay at point *c*, the tangential point between the new budget constraint and the indifference curve I_2 which represents a higher utility level than the one at I_1. However, if the worker is bound by contract to work for *h**, which was determined in Figure 4.1, he is unable to reduce work hours *explicitly*. An alternative way to work less would be to shirk for as much as $L_C - L_S$ as in Figure 4.2. Following the theoretical discussion above, it is predicted that an increase in commuting time would induce workers to work less. Therefore, the testable hypothesis derived is that workers who face longer commuting time provide less effort than otherwise comparable workers.

4.3. Data and methodology

4.3.1. Effort measures

As Engellandt and Riphahn (2005) suggest, it is a widely accepted convention to use absenteeism and unpaid overtime hours as a proxy measure for effort in the related literature. In the studies of determinants for effort, Barmby (2002), Barmby et al. (1991, 1995, 2002), and Johannson and Palme (1996, 2002) measure effort by absenteeism, whilst Booth et al. (2002) and Lazear and Rosen (1990) adopt unpaid overtime hours as a proxy for effort and Engellandt and Riphahn (2005) use both measures. Indeed, absenteeism and unpaid overtime work can be seen as the realisation of unobservable effort at work. Both measures bear one important characteristic of effort: it is beneficial to employers but costly to employees. Furthermore, they are closely related to labour productivity. Van Ommeren and Gutiérrez-i-Puigarnau (2011) use absenteeism as a proxy variable for labour productivity. It is needless to say that workers with lower absenteeism and greater propensity to undertake unpaid overtime work would be more productive than otherwise indistinguishable workers from the employers' point of view.

While both measures appear to be equally well representative of effort in theory and have been treated as such in the empirical literature, there is a practical difference between absenteeism and unpaid overtime work. Absenteeism is discouraged by employers and hence its occurrence is suppressed whilst unpaid overtime work is beneficial to them and thus demanded and encouraged. Workers can do unpaid work only as much as they wish to but they cannot take as many sick-days as they may desire due to the risk of penalisation such as a pay-cut or

dismissal. Another distinctive feature of absenteeism, which differentiates it from unpaid overtime, is its randomness in terms of timing, frequency and duration. In addition to being caused by a lack of underlying work effort, additional causes can be uncertain such as by unpredicted illness or injury and this characteristic of absenteeism further weakens its relationship with underlying effort. In conclusion, absenteeism may not be free from weaknesses as a proxy but this chapter adopts it alongside overtime hours and would interpret the empirical results taking its characteristics into consideration rather than rejecting any of them in advance.

4.3.2. Model specifications

The purpose of the empirical analysis is to test the hypothesis discussed in Section 4.2. Two base-line regression models are specified for this purpose in *eq*. (1) and (2). The absenteeism rate and unpaid overtime hours are expressed as a function of commuting time and other control variables respectively. The model specifications will be estimated by OLS.

$$A_i = \alpha_0 + \alpha_1 C_i + \alpha_n X_i + \varepsilon_i \quad (1)$$

$$U_i = \beta_0 + \beta_1 C_i + \beta_n X_i + \mu_i \quad (2)$$

Where A_i and U_i are the absenteeism rate and unpaid overtime hours, C_i is commuting time, X_i is other controls and ε_i and μ_i are error terms for worker *i*.

As the absenteeism rate is a negative presentation of effort whilst unpaid overtime is a positive one, the former increases (i.e. $\alpha_1 > 0$) and the latter decreases (i.e. $\beta_1 < 0$) with commuting time.

4.3.3. Dataset and variables

The UK Labour Force Survey (LFS) provides measures of both absenteeism and unpaid overtime for individual workers along with other information relevant to the topic of this chapter. It is a quarterly rotating panel survey and has been carried out since 1973 to provide information on socio-demographic characteristics, economic activities and educational characteristics of the UK population, which can be integrated to capture the various trends in the UK labour market. Since 1993, with an increase in surveying frequency from yearly to quarterly, the sample size has been significantly extended. About 120,000 individual respondents come from 15,000 randomly selected households in Great Britain for each quarter, except for the second quarter when the number of households surveyed increases to 40,000 from Great Britain and an additional 4,000 from Northern Ireland.

Unpaid overtime hours are given as the number of hours per week. There are two types of unpaid overtime hours in the QLFS: *usual* and *actual* hours. The actual hours are measured in the reference week (the week previous to when the survey was taken), whilst the usual hours refer to the average of weekly unpaid overtime hours during the quarter of the survey. Although the results are unlikely to change much according to which are selected, usual unpaid overtime hours are chosen to remove the possibility that random events in the reference week create a substantial difference between the actual and usual hours for individual workers.

The QLFS provides several sources for the measurement of absenteeism. It first details whether or not a survey respondent was absent from work in the reference

week. This dichotomous indicator can be used as a dummy dependent variable in the logit or probit models. However, this measure is inferior to continuous measures of absenteeism such as the absenteeism rate. The QLFS reports the numbers of days on which workers were scheduled to work and actual days worked in the reference week. Then, a ratio of the latter to the former can be accepted as an absenteeism rate. However, a 'day' may not be a fine exposure of amount of work especially when there are chances that workers have only part of a full day off. Barmby (2002) suggests a ratio of the number of *hours* taken off to contracted hours as a more sophisticated measure of the absenteeism rate and only acknowledges hours of absenteeism when they are caused by illness, arguing sickness absenteeism may not be caused only by a medical condition. Indeed, the decision to be absent allegedly due to illness is ultimately up to individual workers, although sickness or injury absenteeism does not necessarily result from nefarious motives.

This chapter follows Barmby's (2002) practice regarding the calculation of the absenteeism rate but additional reasons for absenteeism beyond 'sickness and injury' will be considered. In the QLFS, workers who worked fewer hours than usual are asked to state the reasons behind the absences. In most cases, the absences were inevitable, legitimate or pre-acknowledged by employers (e.g. maternity leave, variant work hours, holiday, etc.) but some workers state the reasons they could use to easily excuse themselves for their absences such as bad weather, personal or family-related reasons, and other. In this chapter, all of these cases would be classed under worker's discretionary absenteeism. Then, the absenteeism rate would be expressed as:

$$A_i = \frac{\sum D_i}{\sum C_i} \tag{3}$$

Where A_i is the absenteeism rate, D_i is the number of hours taken off and C_i is the contracted hours of work for worker i.

Then, each component of the absenteeism rate is given as follows:

$$D_i = (h_i^u - h_i^a)s_i \tag{4}$$

$$C_i = h_i^a(1 - s_i) + h_i^u s_i \tag{5}$$

Where h_i^u is usual hours, h_i^a is actual hours and s_i is an absenteeism indicator.

In *eq.*(4), absenteeism is regarded as the difference between usual and actual hours of work. s_i is an indicator for absenteeism and set equal to 1 if the reasons for absence were illness, bad weather, personal/family-related or other, otherwise zero value is assigned. If a worker was not absent at all in the reference week, s_i takes 0 and subsequently D_i and A_i are equal to zero. If the worker took any hours off from work for the reasons listed above, s_i takes the value of 1 and D_i is now equal to the difference between the contracted hours and the actual hours of work (i.e. hours taken off) and C_i is the usual work hours. Then, the absenteeism rate A_i is given as the ratio of D_i and C_i and a positive number. In the QLFS, the usual hours and the actual hours of work are given respectively

by ‘basic usual hours’ and ‘basic actual hours’, both of which are exclusive of any overtime hours.

The daily commuting time is the main explanatory variable and initially given, in the QLFS, as minutes taken to go to work from home but is represented in hours in this chapter for easier interpretation of the estimated coefficients since unpaid overtime work, one of the dependent variables, is also given in hours. Its maximum value is set equal to 3 hours (180 minutes) and any commuting time over 3 hours will be treated as 3 hours. Other explanatory variables are largely categorised into worker’s demographic and other personal characteristics, employment characteristics and regional and year dummies, some of which are discussed in detail below.

A combination of age and age squared is thought to be correlated with both commuting time and effort level as they are highly related to workers’ physical strength which, in turn, can determine how far they can travel to work and how hard they can work. Age can also be an indicator of a worker’s ambition for future career development and so, young workers may exert more effort and commute longer for future promotion opportunities. Gender is also a strong determinant of both commuting time and effort. Female workers usually travel shorter distances to work, work less overtime and take more time off work. This is thought to be related to child and family care demands. Education level is included in addition to basic socio-demographic factors as it is shown in Barmby et al. (2002) to affect absenteeism.

Of the employment conditions, wage is worthy of particular attention as it is positively correlated with both effort and commuting distance empirically. High-income workers are certainly motivated by their wages and hence provide a high level of effort. Barmby et al. (1991) and Barmby (2002) examine the effect of financial incentives on effort and show the potential importance of financial aspects in explaining workers' absenteeism behaviour. At the same time, they are likely to travel a long distance to work partly because high-income workers tend to reside in suburban areas in pursuit of a good residential environment, whereas their workplaces are often located in central business districts. This may also partly be because long-distance commuters are presumed to be compensated by high wage levels (Van Ommeren et al., 1997; Manning, 2003). Overall, high-income workers are characterised by relatively high effort and long commuting time. Therefore, if wage is omitted from the relationships between wage, effort and commuting time, effort would appear positively correlated with commuting time, when they actually have a negative relationship. Of the various types of wages in the QLFS, net weekly wage - the amount of labour income earned weekly after tax - is adopted. The nominal wages from multiple years (2004 to 2010) are deflated by the UK retail price index to the real wage at 2004 prices. The real wage enters the regression model in the form of a dummy to see if it has any non-linear relationship with the dependent variables. See the summary tables below for the detailed categories.

Additionally, a few variables related to workers' job status and characteristics will be included. Engellandt and Riphahn (2005) argue that work motivation levels vary between temporary and permanent workers. Moreover, temporary workers

are likely to have a shorter commuting time than those with permanent positions. Therefore, an indicator of whether a worker is temporary or permanent should be controlled for to reduce potential omitted variable biases. Likewise, it is also expected that part-time and full-time workers differ in terms of hours of work, commuting time and the incentive to put forth effort. Other important variables include type of occupation, as these are considered closely related to not only the nature of work (e.g. usual amount of overtime required, types of duty, etc.) but also commuting patterns and time. There are 9 types of occupation which are listed in the summary statistics tables below.

As the sample is collected between 2004 and 2010, year dummies are used to capture unobservable differences among the years (e.g. unemployment rate and other macro-economic conditions). Lastly, dummies for regions of work are added to capture the regional differences, including the transportation modes and infrastructure. There are two types of regions in the QLFS – regions of work and regions of usual residence. The inclusion of both types of regions would capture the effect of commuting time on the dependent variables to some extent and hence only one of them should be included. Of the two types of regions, the regions of work are selected as they turn out to be more highly correlated with the dependent variables and commuting time. However, the two types of regions coincide in most cases of the regression sample and hence the choice of types of region does not affect the empirical results to any great extent. The QLFS initially recognises 23 different regions across the UK. For simplicity, they are reduced to 12 Government Office Regions. It turns out that the simplification of work regions does not affect the empirical results.

4.3.4. Sample selection

The survey respondents of the QLFS are asked to report their commuting time only in the July-to-September quarter survey. In order to obtain the empirical results from the relatively recent years for policy-relevance, and to maintain consistency for important variables over time, as well as to obtain a reasonable sample size, this chapter collects the data from each July-to-September quarter in the seven years from 2004 to 2010. A pool of total observations from the July-to-September quarters for the seven years amounts to 945,527 cases. However, the information on commuting time is available only for those who worked or did not work in the reference week temporarily but had a job and whose workplaces are separated from home, meaning the size of the potential sample is reduced to 297,793. The number of cases is substantially reduced again to 90,479 as the wage information is available only for those at Wave 1 and 5. Even at the cost of a large sample loss, wage has to be kept in the regression models as it has a correlation with both commuting time and worker effort and therefore its omission could cause bias. When the cases with missing values for the variables appearing in the regression specifications are excluded, the final sample size used for the regressions for absenteeism rate is 78,029. The summary statistics are presented in Table 4.1. For the unpaid overtime regressions, 78,302 cases finally remain after observations with missing values for any variables entering the model are excluded. See Table 4.2 for the summary statistics.

Table 4.1. Summary statistics for absenteeism rate regression

Variable	Obs.	Mean	Std. Dev.	Min	Max
Absenteeism rate	78029	0.033	0.164	0	1
Travel time (in min.)	78029	25.633	22.703	1	180
Age	78029	40.841	12.634	16	84
Age squared	78029	1827.588	1051.837	256	7056
Male (*Female exc.*)	78029	0.466	0.499	0	1
Qualification (1st or higher degree exc.)					
Higher education	78029	0.108	0.311	0	1
A level or equivalent	78029	0.229	0.420	0	1
GCSE or equivalent	78029	0.233	0.423	0	1
Other qualifications	78029	0.112	0.316	0	1
No qualification	78029	0.082	0.275	0	1
Marital status (Single exc.)					
Married	78029	0.553	0.497	0	1
Separated	78029	0.030	0.170	0	1
Divorced	78029	0.086	0.280	0	1
Widowed	78029	0.014	0.116	0	1
Dep. children in family	78029	0.376	0.484	0	1
Sex & dependent children (Male & no dep. child exc.)					
Male & dep. child	78029	0.170	0.375	0	1
Female & no dep. child	78029	0.328	0.469	0	1
Female & dep. child	78029	0.206	0.404	0	1
Real wage band (Less than £100 exc.)					
£100~£199	78029	0.246	0.431	0	1
£200~£299	78029	0.268	0.443	0	1
£300~£399	78029	0.166	0.372	0	1
£400~£499	78029	0.092	0.289	0	1
£500 or over	78029	0.093	0.291	0	1
Full-time worker (Part-time exc.)	78029	0.724	0.447	0	1
Permanent worker (Temp exc.)	78029	0.950	0.218	0	1

Table 4.1. Summary statistics for absenteeism rate regression (cont.)

Variable	Obs.	Mean	Std. Dev.	Min	Max
Job tenure (Less than 1 yr exc.)					
1 but less than 5 years	78029	0.334	0.472	0	1
5 but less than 10 years	78029	0.208	0.406	0	1
10 less than 20 years	78029	0.184	0.387	0	1
20 years or more	78029	0.115	0.318	0	1
Occupational class (Managers & senior official exc.)					
Professional occ.	78029	0.136	0.343	0	1
Assoc. profession &technical	78029	0.147	0.354	0	1
Administrative and secretarial	78029	0.137	0.344	0	1
Skilled trades occupations	78029	0.072	0.259	0	1
Personal service occupations	78029	0.090	0.287	0	1
Sales and customer service	78029	0.083	0.277	0	1
Process, plant and machine	78029	0.070	0.256	0	1
Elementary occupations	78029	0.126	0.332	0	1
Usual weekly work hours (0~14 hours exc.)					
15~24 hours	78029	0.149	0.356	0	1
25~34 hours	78029	0.108	0.311	0	1
35 hours or over	78029	0.670	0.470	0	1
Regional dummies (North East exc.)					
North West	78029	0.104	0.305	0	1
Yorkshire & the Humber	78029	0.098	0.298	0	1
East Midlands	78029	0.078	0.268	0	1
West Midlands	78029	0.084	0.278	0	1
Eastern	78029	0.039	0.193	0	1
London	78029	0.105	0.307	0	1
South East	78029	0.186	0.389	0	1
South West	78029	0.086	0.281	0	1
Wales	78029	0.045	0.208	0	1
Scotland	78029	0.095	0.293	0	1
Northern Ireland	78029	0.023	0.149	0	1
Year (2004 exc.)					
2005	78029	0.114	0.317	0	1
2006	78029	0.164	0.370	0	1
2007	78029	0.164	0.370	0	1
2008	78029	0.159	0.366	0	1
2009	78029	0.144	0.351	0	1
2010	78029	0.137	0.343	0	1

Table 4.2. Summary statistics for unpaid overtime work regression

Variable	Obs.	Mean	Std. Dev.	Min	Max
Unpaid overtime hours	78302	1.520	4.002	0	79
Travel time (in min.)	78302	25.664	22.792	1	180
Age	78302	40.850	12.650	16	84
Age squared	78302	1828.719	1053.443	256	7056
Male (*Female exc.*)	78302	0.467	0.499	0	1
Qualification (1st or higher degree exc.)					
Higher education	78302	0.108	0.311	0	1
A level or equivalent	78302	0.229	0.420	0	1
GCSE or equivalent	78302	0.233	0.423	0	1
Other qualifications	78302	0.112	0.316	0	1
No qualification	78302	0.082	0.275	0	1
Marital status (Single exc.)					
Married	78302	0.553	0.497	0	1
Separated	78302	0.030	0.170	0	1
Divorced	78302	0.086	0.280	0	1
Widowed	78302	0.014	0.117	0	1
Dep. children in family	78302	0.376	0.484	0	1
Sex & dependent children (Male & no dep. child exc.)					
Male & dep. child	78302	0.170	0.376	0	1
Female & no dep. child	78302	0.328	0.469	0	1
Female & dep. child	78302	0.205	0.404	0	1
Real wage band (Less than £100 exc.)					
£100~£199	78302	0.246	0.430	0	1
£200~£299	78302	0.268	0.443	0	1
£300~£399	78302	0.166	0.372	0	1
£400~£499	78302	0.092	0.289	0	1
£500 or over	78302	0.093	0.291	0	1
Full-time worker (Part-time exc.)	78302	0.723	0.447	0	1
Permanent worker (Temp exc.)	78302	0.949	0.219	0	1

Table 4.2. Summary statistics for unpaid overtime work regression (cont.)

Variable	Obs.	Mean	Std. Dev.	Min	Max
Job tenure (Less than 1 yr exc.)					
1 but less than 5 years	78302	0.334	0.472	0	1
5 but less than 10 years	78302	0.208	0.406	0	1
10 less than 20 years	78302	0.184	0.387	0	1
20 years or more	78302	0.114	0.318	0	1
Occupational class (Managers & senior official exc.)					
Professional occ.	78302	0.136	0.343	0	1
Assoc. profession &technical	78302	0.147	0.354	0	1
Administrative and secretarial	78302	0.137	0.343	0	1
Skilled trades occupations	78302	0.072	0.259	0	1
Personal service occupations	78302	0.090	0.287	0	1
Sales and customer service	78302	0.083	0.276	0	1
Process, plant and machine	78302	0.070	0.256	0	1
Elementary occupations	78302	0.126	0.332	0	1
Paid overtime work hours	78302	1.249	3.503	0	60
Usual weekly work *hours* (0~14 hours exc.)					
15~24 hours	78302	0.148	0.355	0	1
25~34 hours	78302	0.108	0.310	0	1
35 hours or over	78302	0.671	0.470	0	1
Regional dummies (North East exc.)					
North West	78302	0.104	0.305	0	1
Yorkshire & the Humber	78302	0.098	0.298	0	1
East Midlands	78302	0.078	0.268	0	1
West Midlands	78302	0.084	0.278	0	1
Eastern	78302	0.039	0.193	0	1
London	78302	0.105	0.307	0	1
South East	78302	0.186	0.389	0	1
South West	78302	0.087	0.281	0	1
Wales	78302	0.045	0.208	0	1
Scotland	78302	0.095	0.293	0	1
Northern Ireland	78302	0.023	0.148	0	1
Year (2004 exc.)					
2005	78302	0.114	0.317	0	1
2006	78302	0.164	0.370	0	1
2007	78302	0.164	0.370	0	1
2008	78302	0.159	0.366	0	1
2009	78302	0.143	0.351	0	1
2010	78302	0.137	0.343	0	1

4.4. Empirical results and discussion

4.4.1. Absenteeism rate

Table 4.3 reports the OLS regression results for the absenteeism rates. The explanatory variables are gradually added from column (1) to column (6). This is done so it would be possible to see how explanatory variables are correlated with commuting time and absenteeism rates and if they reduce bias. In column (1), only commuting time is controlled for and it turns out to have a negative correlation with the absenteeism rate implying that workers who commute longer tend to exhibit lower degree of absenteeism, which is contrary to the prediction made in Section 4.2. However, the correlation is not significant at any conventional levels. In column (2), the addition of regions of work and year dummies increases the size of the estimate of commuting time slightly and the p-value also decreases towards zero (i.e. the estimate becomes more significant).

The job characteristics are controlled for first as they are believed to be more highly correlated with commuting time and the absenteeism rate than personal and demographic attributes. Among the job characteristics, wage deserves prior attention as it is expected to be the most relevant in explaining the spurious relationship between the dependent and the main explanatory variables. When the dummy variables for the real wage bands are controlled for in column (3), the estimated coefficient of commuting time increases dramatically and is now statistically significant at 1% level. As predicted previously, the labour income turns out to be correlated positively with both commuting time and work effort. Furthermore, judging from the estimates of the real wage dummies, the absenteeism rate goes down as the wage increases. This result is consistent with

the findings in the related literature, which suggests that the financial rewards exert a positive influence over work effort levels (Barmby, 2002). In column (4), the two variables which describe personal job status within the workers' firms are added. They indicate respectively whether workers are full-time or part-time and whether temporary or permanent. Contrary to the predictions made previously, the estimated coefficient of interest changes little. This result implies that such a status does not have an independent impact on the estimate, separate from the wage level.

When the remaining job characteristics are added in column (5), the estimated coefficient of commuting time increases and becomes more statistically significant according to the attached p-value. The job tenure (the length of time in the current job) turns out to be important in uncovering the spurious relationship between commuting time and the absenteeism rate. The job tenure is found to be positively correlated with the absenteeism rate. Those who have worked for their current employers for less than 1 year have an absenteeism rate of about 2.52%, whereas those there for over 20 years have a 3.76% absenteeism rate. Furthermore, the estimated coefficients of the job tenure dummies suggest that workers tend to take more time off due to sickness and injury and personal- or family-related reasons as the time with their employer increases. This could partly be for reasons of age and health but it rather seems because workers become more comfortable with revealing their desire for absenteeism as the time spent in the same workplace increases. Indeed, the addition of age and age squared does not change the magnitude of the estimates of job tenure a great deal, as in column (6), which is supportive of this interpretation. Job tenure is correlated negatively with

commuting time. This is likely because workers have gradually moved closer to their workplaces over time (Gordon, Richardson and Jun, 1991). Therefore, the addition of job tenure corrects the negative bias and increases the estimate of main interest.

The occupational type or class also increases the estimate slightly. Of the occupations, workers in high-status careers are found to be absent from work less often and commute for a longer time. While the basic average absenteeism rate for the entire sample is 3.30%, those for high-status jobs such as the ‘managers and senior officials’ and ‘professional occupation’ are 1.93% and 2.71% respectively. The average commuting times for these two groups are 33.27 and 30.91 minutes respectively but the average for all workers is only 25.63 minutes. In contrast, ‘personal service occupation’, which is a relatively low-status job, has the highest absenteeism rate (4.85%) and the second shortest commuting time (18.98 minutes). Those in ‘process, plant and machine operative’ and ‘elementary occupations’ are also characterised by high absenteeism rates and short commuting times. Overall, absenteeism rates and commuting times are negatively correlated across the occupational types and hence their omission could cause downward bias. It is worth noting that its effect on the relationship between the commuting time and the absenteeism rate is independent of the wage. This result implies that high-status jobs may provide other rewards, in addition to wages and therefore a worker’s occupational type seems to indicate a level of non-wage compensation to some extent.

Table 4.3. OLS estimation of absenteeism rate
(Dependent variable: Absenteeism rate)

	(1)	(2)	(3)	(4)	(5)	(6)
Commuting time (in hour)	-0.000250	0.000651	0.00621***	0.00622***	0.00820***	0.00854***
	(0.00163)	(0.00173)	(0.00181)	(0.00181)	(0.00183)	(0.00183)
Real wage band						
(Less than £100 exc.)						
£100~£199			-0.005**	-0.006**	-0.010***	-0.011***
			(0.002)	(0.003)	(0.003)	(0.003)
£200~£299			-0.010***	-0.013***	-0.017***	-0.018***
			(0.002)	(0.003)	(0.003)	(0.003)
£300~£399			-0.014***	-0.017***	-0.022***	-0.022***
			(0.002)	(0.003)	(0.003)	(0.004)
£400~£499			-0.022***	-0.025***	-0.029***	-0.029***
			(0.003)	(0.003)	(0.004)	(0.004)
£500 or over			-0.026***	-0.028***	-0.031***	-0.031***
			(0.002)	(0.003)	(0.004)	(0.004)
Full-time worker (Part-time exc.)				0.003	0.005***	0.008***
				(0.002)	(0.002)	(0.002)
Permanent worker (Temp exc.)				0.003	-0.003	-0.004
				(0.003)	(0.003)	(0.003)
Job tenure						
(Less than 1 yr exc.)						
1 but less than 5 years					0.010***	0.008***
					(0.002)	(0.002)
5 but less than 10 years					0.016***	0.012***
					(0.002)	(0.002)
10 less than 20 years					0.018***	0.012***
					(0.002)	(0.002)
20 years or more					0.021***	0.015***
					(0.002)	(0.003)

Table 4.3. OLS estimation of absenteeism rate (cont.)

	(1)	(2)	(3)	(4)	(5)	(6)
Occupational class						
(Managers & senior official exc.)						
Professional occ.					0.009***	0.008***
					(0.002)	(0.002)
Assoc. profession & technical					0.014***	0.014***
					(0.002)	(0.002)
Administrative and secretarial					0.011***	0.010***
					(0.002)	(0.002)
Skilled trades occupations					0.007***	0.010***
					(0.002)	(0.003)
Personal service occupations					0.023***	0.022***
					(0.003)	(0.003)
Sales and customer service					0.007***	0.009***
					(0.003)	(0.003)
Process, plant and machine					0.014***	0.014***
					(0.003)	(0.003)
Elementary occupations					0.011***	0.012***
					(0.002)	(0.003)
Age						0.001**
						(0.000)
Age squared						-0.000
						(0.000)
Male (Female exc.)						-0.005***
						(0.001)
Qualification						
(1st or higher degree exc.)						
Higher education						-0.004*
						(0.002)

Table 4.3. OLS estimation of absenteeism rate (cont.)

	(1)	(2)	(3)	(4)	(5)	(6)
A level or equivalent						-0.003
						(0.002)
GCSE or equivalent						-0.001
						(0.002)
Other qualifications						-0.002
						(0.002)
No qualification						0.001
						(0.003)
Marital status (Single exc.)						
Married						0.001
						(0.002)
Separated						0.000
						(0.004)
Divorced						0.007**
						(0.003)
Widowed						0.012*
						(0.007)
Dep. children in family						-0.000
						(0.001)
Region	No	Yes	Yes	Yes	Yes	Yes
Year	No	Yes	Yes	Yes	Yes	Yes
Constant	0.033***	0.034***	0.041***	0.039***	0.022***	0.007
	(0.001)	(0.003)	(0.003)	(0.004)	(0.005)	(0.008)
Adj. R-squared	-0.000	0.000	0.002	0.002	0.004	0.005
No. of obs.	78029	78029	78029	78029	78029	78029

Note: ***, **,* refer to the significance levels at 1, 5 and 10% respectively. Robust standard errors are given in parenthesis. The explanatory variables are added to the specification gradually. The types of variables controlled for additionally for each column are as follows: (1) commuting time (2) regional and year dummies (3) real wage (4) personal job statuses (5) other job characteristics (6) personal characteristics. The estimated coefficients for the regional and year dummies are missed from the table.

The inclusion of personal and demographic characteristics does not change the estimate to a great degree, as shown in column (6). Of the personal characteristics, gender is the most relevant in explaining the spurious relationship and in fact, it is almost entirely responsible for the increase in the estimate. The reason for the increased estimate of commuting time is that gender is correlated with both commuting time and absenteeism. As shown with the negative sign for male workers in column (4), female workers have higher absenteeism rates than male workers on average. In the literature, women have been also reported to have a higher rate of absenteeism than men (Johns, 1978; Garcia, 1987; Barmby et al., 2002). A perception that women are physically more vulnerable could be a reason for which absenteeism is allowed more generously for women. Mastekaasa and Olsen (1998) argue that a gender difference in absenteeism reflects a difference in general health conditions and personality between the two genders. At the same time, commuting time is longer for men on average (28.88 vs. 22.80 minutes). Without including gender in the specifications, then, shorter-distance commuters would appear to be more often absent from work and the estimate of commuting time would be biased downward.

The regression result in column (6), with all the relevant independent variables being controlled for, suggests that there exists a positive correlation between commuting time and the absenteeism rate, which is consistent with the prediction. As sensed from the low R squared, only a little of the variation in the absenteeism can be explained by control variables, probably because, by nature, absenteeism is often caused by *random* events and is suppressed by employers. Considering this, one of the contributions made by this chapter is that commuting time is identified

as one of the few strong determinants of absenteeism. An hour increase in commuting time leads to a 0.85% increase in the absenteeism rate. Considering that the average absenteeism rate is 3.3%, as shown in Table 4.1, this increase does not seem trivial.

4.4.2. Unpaid overtime hours

Table 4.4 reports the OLS estimation results for usual unpaid overtime hours. Similar to the case of absenteeism, the model specifications vary from (1) to (6) by adding more control variables gradually. In column (1), where only the daily commuting time is included, the estimate of commuting time is highly positive and statistically very significant, which is contrary to the prediction made in Section 4.2. However, the estimate is likely to be biased due to many omitted variables. Even when, in addition, the regional and year dummies are controlled for, the bias slightly decreases (column (2)). In column (3), the real wage dummies are included first among the various job characteristics. The estimate is substantially reduced and turns negative, which is consistent with the prediction made previously but it is barely significant yet. Unpaid overtime work increases with the wage non-linearly, in particular, at an increasing rate. Those who earn over £500 per week do 4.8 hours more of unpaid work per week than those who earn less than £100. Reflecting the case of absenteeism, personal job status within companies has little independent effect on the estimate from the real wage, when those three variables are controlled for in column (4).

Table 4.4. OLS estimation of unpaid overtime work
(Dependent variable: Unpaid overtime work hours)

	(1)	(2)	(3)	(4)	(5)	(6)
Commuting time (in hour)	1.154***	0.994***	-0.0503	-0.0526	-0.170***	-0.176***
	(0.0470)	(0.0489)	(0.0479)	(0.0480)	(0.0469)	(0.0465)
Real wage band						
(Less than £100 exc.)						
£100~£199			0.154***	0.049*	-0.144***	-0.222***
			(0.019)	(0.025)	(0.026)	(0.028)
£200~£299			0.778***	0.589***	0.088**	-0.046
			(0.027)	(0.037)	(0.040)	(0.042)
£300~£399			1.827***	1.625***	0.635***	0.444***
			(0.042)	(0.050)	(0.055)	(0.058)
£400~£499			3.249***	3.041***	1.539***	1.296***
			(0.070)	(0.075)	(0.079)	(0.083)
£500 or over			4.808***	4.598***	2.772***	2.470***
			(0.083)	(0.087)	(0.094)	(0.097)
Full-time worker				0.223***	0.489***	0.620***
				(0.026)	(0.027)	(0.030)
Permanent worker				0.102**	0.290***	0.346***
				(0.049)	(0.049)	(0.050)
Job tenure						
(Less than 1 yr exc.)						
1 but less than 5 years					0.170***	0.163***
					(0.035)	(0.036)
5 but less than 10 years					0.131***	0.142***
					(0.041)	(0.043)
10 less than 20 years					0.036	0.086*
					(0.044)	(0.047)
20 years or more					-0.023	0.098*
					(0.053)	(0.057)

Table 4.4. OLS estimation of unpaid overtime work (cont.)

	(1)	(2)	(3)	(4)	(5)	(6)
Occupational class						
(Managers & senior official exc.)						
Professional occ.					0.432***	0.042
					(0.081)	(0.081)
Assoc. profession & technical					-1.731***	-1.831***
					(0.062)	(0.062)
Administrative and secretarial					-1.985***	-1.931***
					(0.059)	(0.059)
Skilled trades occupations					-2.477***	-2.089***
					(0.061)	(0.063)
Personal service occupations					-1.880***	-1.810***
					(0.061)	(0.061)
Sales and customer service					-2.103***	-1.901***
					(0.060)	(0.061)
Process, plant and machine					-2.495***	-2.076***
					(0.061)	(0.063)
Elementary occupations					-2.285***	-1.942***
					(0.057)	(0.058)
Paid overtime work hours					-0.076***	-0.068***
					(0.003)	(0.003)
Age						0.023***
						(0.006)
Age squared						-0.000***
						(0.000)
Male (Female exc.)						-0.332***
						(0.035)
Qualification						
(1st or higher degree exc.)						
Higher education						-0.630***
						(0.060)

Table 4.4. OLS estimation of unpaid overtime work (cont.)

	(1)	(2)	(3)	(4)	(5)	(6)
A level or equivalent						-0.960***
						(0.048)
GCSE or equivalent						-0.997***
						(0.047)
Other qualifications						-1.113***
						(0.052)
No qualification						-1.213***
						(0.050)
Marital status (Single exc.)						
Married						0.037
						(0.038)
Separated						-0.041
						(0.072)
Divorced						0.098*
						(0.056)
Widowed						-0.017
						(0.095)
Dep. children in family						-0.116***
						(0.031)
Region	No	Yes	Yes	Yes	Yes	Yes
Year	No	Yes	Yes	Yes	Yes	Yes
Constant	1.026***	0.837***	0.116*	0.002	1.911***	2.299***
	(0.022)	(0.070)	(0.065)	(0.079)	(0.096)	(0.135)
Adj. R-squared	0.012	0.016	0.136	0.136	0.197	0.206
No. of obs.	78302	78302	78302	78302	78302	78302

Note: ***, **,* refer to the significance levels at 1, 5 and 10% respectively. Robust standard errors are given in parenthesis. The explanatory variables are added to the specification gradually. The types of variables controlled for additionally for each column are as follows: (1) commuting time (2) regional and year dummies (3) real wage (4) personal job statuses (5) other job characteristics (6) personal characteristics. The estimated coefficients for the regional and year dummies are missed from the table.

When the remaining job characteristics are controlled for, the estimate deceases substantially and becomes statistically significant at 1% level (column (5)). Among the newly added variables in column (5), the types of occupation are particularly effective in correcting the bias of the estimated coefficient of commuting time. Those with the two highest-status jobs (manager and senior official and professional occupations) do more unpaid work per week by 3 to 4 hours compared to the rest. Also, their commuting time is well over 30 minutes while the average is only 26 minutes. On the other hand, those with relatively low skilled jobs are likely to be characterised by low unpaid overtime work (usually less than an hour) and short commuting time (about 20 minutes). As was the case with absenteeism, the occupational types seem to be deeply associated with the non-wage compensations and therefore their inclusion removes the bias further.

Paid overtime hours appear only in the regression specification for the unpaid overtime. When being controlled for in column (5), it decreases the estimate of interest, implying that commuting time and unpaid overtime hours were spuriously correlated through the paid overtime hours. Firstly, the paid overtime work has a strong negative relationship with the amount of unpaid overtime work, likely because there is a time constraint on how much more workers can work on top of their basic hours. For the regression sample, the total overtime work (a sum of unpaid and paid overtime work) averages at just over 6 hours and 75% of the workers work overtime less than 10 hours and 90% less than 15 hours per week. That is, a majority of workers can work only up to certain limited hours per week. Therefore, there has to be a trade-off between paid and unpaid overtime hours from the workers' points of view. Indeed, among the regression sample, those who

did less than an hour-long *paid* overtime did over 5 hours of *unpaid* overtime work, whereas those who did more than 5 hours *paid* overtime only did less than an hour of *unpaid* work. As an indicator for the degree of the trade-off, the correlation between them is -0.327. At the same time, workers with shorter commuting time tend to work more paid overtime than those with longer commuting time. This seems because those with short commuting time are less time-constrained and hence they are able to do more work to earn more. As a result, if the paid overtime hours were not included in the model, the estimated coefficient of commuting time would be biased upward.

In column (6), the respondents' personal and demographic characteristics are added and the estimate becomes more negative. Unlike in the case of absenteeism, regarding personal characteristics, educational qualification level plays the biggest role in correcting the downward bias. This is because the level of education is positively correlated with both commuting time and unpaid overtime hours. It is relatively well reported that educational attainment is closely associated with commuting time (e.g. Schwanen and Dijst, 2002; Lee and McDonald, 2003). However, it is unclear why the education level is found to be correlated with unpaid overtime work in a way that highly educated workers tend to do more unpaid hours. The best guess is that those with a higher level of education are more ambitious for their future career path so that by doing more unpaid work they show their loyalty to the employers and acquire work skills and experience. The indirect evidence found in the literature suggests that long-term unpaid work increases real labour earnings and the probability of promotion, or job retention (Anger, 2006; Pannenberg, 2002). Furthermore, it may be partly because jobs

wanted by highly educated workers are relatively rare and pay a wage high enough to attract workers from long distances.

Despite the absenteeism rate and unpaid overtime work hours being fairly distinguishable in terms of their characteristics, except that they are both proxies for effort level, the two sets of empirical results which adopt each of them as a dependent variable respectively indicate the same findings. These findings, moreover, are consistent with the predictions made in Section 4.3. As most major variables are controlled for, the likelihood of the bias caused by omitted variables should be fairly low. Furthermore, on the whole, the addition of explanatory variables tends to reduce bias.

4.4.3. Regressions with interaction terms

When various personal, demographic and job characteristics are controlled for, the relationship between work effort and commuting time turns out to be statistically significant. One way to check the robustness of the results is to see if the empirical results are consistent with predictions for sub-sample groups using interaction terms. Firstly, the estimated coefficient of commuting time is examined by the full-time/part-time status of workers. In essence, the neoclassical model predicts that commuting time is considered as a part of work hours but not paid for so that the *effective* hourly wage (income divided by a sum of working and commuting hours) goes down as commuting time increases. As a result, the work effort level is also expected to become lower as the effective wage declines. An increase in commuting time would lower the effective wage for part-timers to a greater extent than for full-timers and therefore the effort level would fall more

sharply with commuting time for part-timers. The estimated coefficient of commuting time in the absenteeism regression should be greater and the one in the unpaid overtime work regression should be lower (more negative) for part-time workers than for full-time workers.

The regression results with the interaction terms between the full-time/part-time indicators and commuting time are presented in Table 4.5 and column (1) shows the results for the absenteeism regression. For both part-time and full-time workers, the estimated coefficients of commuting time are positive and statistically significant. This shows that the relationship is not confined just to a certain sub-sample group. More importantly, the estimated coefficient is much greater for part-time workers than full-time workers as expected. Column (2) presents the results for the unpaid overtime regression. Similarly, the estimated coefficients are negative in both cases with that for part-time workers being more pronounced. Again, the empirical results are consistent with the prediction regardless of the type of dependent variables.

Table 4.5. OLS regression with interaction terms 1

Dependent variable	(1) Absenteeism rate	(2) Unpaid overtime	(3) Absenteeism rate	(4) Unpaid overtime
Interaction terms:				
Full-time × commuting time	0.00644***	-0.111**		
	(0.00194)	(0.0539)		
Part-time × commuting time	0.0193***	-0.503***		
	(0.00472)	(0.0507)		
Interaction terms:				
0~14 hrs × commuting time			0.0262**	-0.506***
			(0.0108)	(0.0752)
15~24 hrs × commuting time			0.0150***	-0.504***
			(0.00568)	(0.0795)
25~34 hrs × commuting time			0.0147**	-0.377***
			(0.00604)	(0.129)
35 hrs and over ×commuting time			0.00637***	-0.0913*
			(0.00200)	(0.0551)
Real wage band (Less than £100 exc.)				
£100~£199	-0.012***	-0.196***	-0.015***	-0.157***
	(0.003)	(0.028)	(0.003)	(0.036)
£200~£299	-0.019***	-0.016	-0.022***	0.193***
	(0.003)	(0.042)	(0.004)	(0.050)
£300~£399	-0.023***	0.467***	-0.026***	0.707***
	(0.004)	(0.058)	(0.004)	(0.065)
£400~£499	-0.030***	1.312***	-0.033***	1.575***
	(0.004)	(0.082)	(0.004)	(0.088)
£500 or over	-0.031***	2.475***	-0.034***	2.745***
	(0.004)	(0.097)	(0.004)	(0.101)
Full-time worker (Part-time exc.)	0.013***	0.474***		
	(0.003)	(0.040)		
Permanent worker (Temp exc.)	-0.003	0.339***	-0.004	0.375***
	(0.003)	(0.050)	(0.003)	(0.050)
Job tenure (Less than 1 yr exc.)				
1 but less than 5 years	0.008***	0.163***	0.008***	0.154***
	(0.002)	(0.036)	(0.002)	(0.036)
5 but less than 10 years	0.012***	0.144***	0.012***	0.120***
	(0.002)	(0.043)	(0.002)	(0.043)
10 less than 20 years	0.012***	0.089*	0.012***	0.066
	(0.002)	(0.047)	(0.002)	(0.047)
20 years or more	0.015***	0.105*	0.015***	0.082
	(0.003)	(0.057)	(0.003)	(0.057)
Occupational class (Managers & senior official exc.)				
Professional occ.	0.008***	0.044	0.009***	0.013
	(0.002)	(0.081)	(0.002)	(0.081)
Assoc. profession & technical	0.015***	-1.833***	0.015***	-1.862***
	(0.002)	(0.062)	(0.002)	(0.062)
Administrative and secretarial	0.010***	-1.934***	0.010***	-1.953***
	(0.002)	(0.059)	(0.002)	(0.059)
Skilled trades occupations	0.010***	-2.090***	0.010***	-2.083***
	(0.003)	(0.063)	(0.003)	(0.064)
Personal service occupations	0.022***	-1.817***	0.022***	-1.859***
	(0.003)	(0.061)	(0.003)	(0.061)

Table 4.5. OLS regression with interaction terms 1 (cont.)

Dependent variable	(1) Absenteeism rate	(2) Unpaid overtime	(3) Absenteeism rate	(4) Unpaid overtime
Sales and customer service	0.009***	-1.906***	0.009***	-1.943***
	(0.003)	(0.061)	(0.003)	(0.061)
Process, plant and machine	0.014***	-2.075***	0.014***	-2.059***
	(0.003)	(0.063)	(0.003)	(0.064)
Elementary occupations	0.012***	-1.950***	0.012***	-1.975***
	(0.003)	(0.058)	(0.003)	(0.058)
Age	0.001**	0.023***	0.001*	0.023***
	(0.000)	(0.006)	(0.000)	(0.006)
Age squared	-0.000	-0.000***	-0.000	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Male (Female exc.)	-0.005***	-0.332***	-0.005***	-0.251***
	(0.001)	(0.035)	(0.001)	(0.035)
Qualification (1st or higher degree exc.)				
Higher education	-0.004*	-0.632***	-0.004*	-0.633***
	(0.002)	(0.060)	(0.002)	(0.060)
A level or equivalent	-0.002	-0.964***	-0.003	-0.949***
	(0.002)	(0.048)	(0.002)	(0.048)
GCSE or equivalent	-0.001	-1.003***	-0.001	-0.981***
	(0.002)	(0.047)	(0.002)	(0.047)
Other qualifications	-0.002	-1.116***	-0.002	-1.081***
	(0.002)	(0.052)	(0.002)	(0.052)
No qualification	0.001	-1.217***	0.001	-1.174***
	(0.003)	(0.050)	(0.003)	(0.050)
Marital status (Single exc.)				
Married	0.001	0.035	0.001	0.014
	(0.002)	(0.038)	(0.002)	(0.038)
Separated	-0.000	-0.038	-0.000	-0.046
	(0.004)	(0.072)	(0.004)	(0.072)
Divorced	0.007**	0.097*	0.007**	0.097*
	(0.003)	(0.056)	(0.003)	(0.056)
Widowed	0.013*	-0.021	0.013*	-0.022
	(0.007)	(0.095)	(0.007)	(0.096)
Dep. children in family	-0.000	-0.115***	-0.000	-0.170***
	(0.001)	(0.031)	(0.001)	(0.031)
Paid overtime work hours		-0.068***		-0.067***
		(0.003)		(0.003)
Usual weekly work hours (0~14 hours exc.)				
15~24 hours			0.008*	0.074*
			(0.005)	(0.040)
25~34 hours			0.015***	0.561***
			(0.005)	(0.069)
35 hours or over			0.020***	0.179***
			(0.005)	(0.057)
Constant	0.003	2.411***	0.000	2.435***
	(0.008)	(0.135)	(0.008)	(0.136)
Region	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
Adj. R-squared	0.005	0.206	0.005	0.205
No. of obs.	78029	78302	78029	78302

Note: ***, **,* refer to the significance levels at 1, 5 and 10% respectively. Robust standard errors are given in parenthesis. The estimated coefficients for the regional and year dummies are missed from the table.

There could be a concern that part-time/full-time status may represent something other than the length of work hours. The differences between part-time and full-time workers may come from the difference in the degree of commitment to their employers or desire and expectation for future career development within their companies. Depending on the assumption taken regarding what part-time/full-time status really represents, the empirical results above may be interpreted in different ways. To avoid ambiguity in the interpretation of the results, a direct measure of work hours is introduced to replace the full-time/part-time status. Basic work hours in the form of dummy variables are now included in the regression model instead of the full-time/part-time status. The basic work hours are categorised into 0~14, 15~24, 25~34, and 35 hour or over. Then, the estimated coefficients are examined for the four groups of working hours using the interaction terms. Column (3) and (4) of Table 4.5 show respectively the empirical results with the interaction terms between commuting time and the bands of basic work hours. In both columns, it is clearly shown that the relationship between commuting and the dependent variables becomes stronger and statistically more significant as the basic work hours decrease. Therefore, it is confirmed that the length of work hours has an effect on the extent to which commuting time affects work effort, which is consistent with the prediction based on the theoretical model.

It is also the case that the behaviour of men and women regarding commuting and absenteeism or unpaid overtime is very different. In the cases of both unpaid overtime and absenteeism, female workers are found to be associated with a lower level of work effort according to the estimated coefficient of ‘Male’ in Table 4.3 and Table 4.4. This is plausible since women usually do more household work and

hence are more time-constrained than men, especially when they have dependent children. The time-constrained workers with less leisure time would respond more sensitively to a given increase in commuting time so that they would reduce their work effort to a greater extent. Therefore, it is expected that working women are absent more often from work and do less unpaid overtime work than working men as commuting time increases. The estimated coefficient of commuting time would be greater for women in the absenteeism regression and smaller in the unpaid overtime work regression. This prediction is tested using the interaction term between commuting time and gender of the workers. The estimated coefficient for women is greater than that for men in the absenteeism regression (column (1) of Table 4.6) and more negative in the unpaid overtime work regression (column (2) of Table 4.6).

Tests are made to determine whether the presence of dependent children is the main reason why female workers reduce their effort levels to a greater extent than male counterparts as commuting time increases. For this experiment, a new variable is created by combining the sex and the existence of dependent children in the family. The new variable takes four options: ‘male & no dependent children’, ‘male & some dependent children’, ‘female & no dependent children’ or ‘female & some dependent children’. Then, the estimate of commuting time for each of those sub-groups is examined through the interaction terms between the ‘sex & dependent children’ variable and commuting time. Among those four sub-groups, ‘female & some dependent children’ is expected to exhibit the strongest relationship between commuting time and the dependent variables.

Table 4.6. OLS regressions with interaction terms 2

Dependent variable	(1) Absence rate	(2) Unpaid overtime	(3) Absence rate	(4) Unpaid overtime
Interaction terms:				
Sex × Commuting time				
Male × commuting time	0.00652***	-0.116*		
	(0.00222)	(0.0621)		
Female × commuting time	0.0117***	-0.270***		
	(0.00285)	(0.0631)		
Interaction terms:				
Sex & Dep. child status				
×Commuting time				
Male & no dep. Child			0.00868***	-0.165**
×commuting time				
			(0.00291)	(0.0763)
Male &dep. child			0.00318	-0.0442
×commuting time				
			(0.00318)	(0.101)
Female &no dep. Child			0.0107***	-0.193**
×commuting time				
			(0.00355)	(0.0802)
Female & dep. Child			0.0134***	-0.417***
×commuting time				
			(0.00443)	(0.0937)
Real wage band (below £100 exc.)				
£100~£199	-0.011***	-0.217***	-0.011***	-0.210***
	(0.003)	(0.028)	(0.003)	(0.028)
£200~£299	-0.018***	-0.036	-0.018***	-0.033
	(0.003)	(0.043)	(0.003)	(0.043)
£300~£399	-0.022***	0.453***	-0.022***	0.453***
	(0.004)	(0.058)	(0.004)	(0.058)
£400~£499	-0.029***	1.302***	-0.030***	1.298***
	(0.004)	(0.083)	(0.004)	(0.083)
£500 or over	-0.031***	2.470***	-0.031***	2.459***
	(0.004)	(0.097)	(0.004)	(0.097)
Full-time worker (Part-time exc.)	0.008***	0.621***	0.008***	0.599***
	(0.002)	(0.030)	(0.002)	(0.030)
Permanent worker (Temp exc.)	-0.004	0.344***	-0.003	0.346***
	(0.003)	(0.050)	(0.003)	(0.050)
Job tenure (Less than 1 yr exc.)				
1 but less than 5 years	0.008***	0.162***	0.008***	0.163***
	(0.002)	(0.036)	(0.002)	(0.036)
5 but less than 10 years	0.012***	0.142***	0.012***	0.145***
	(0.002)	(0.043)	(0.002)	(0.043)
10 less than 20 years	0.012***	0.086*	0.012***	0.087*
	(0.002)	(0.047)	(0.002)	(0.047)
20 years or more	0.015***	0.099*	0.015***	0.105*
	(0.003)	(0.057)	(0.003)	(0.057)
Occupational class				
(Managers & senior official exc.)				
Professional occ.	0.008***	0.041	0.008***	0.045
	(0.002)	(0.081)	(0.002)	(0.081)
Assoc. profession & technical	0.015***	-1.832***	0.015***	-1.830***
	(0.002)	(0.062)	(0.002)	(0.062)
Administrative and secretarial	0.010***	-1.932***	0.010***	-1.935***
	(0.002)	(0.059)	(0.002)	(0.059)
Skilled trades occupations	0.010***	-2.087***	0.010***	-2.088***
	(0.003)	(0.063)	(0.003)	(0.063)

Table 4.6. OLS regressions with interaction terms 2 (cont.)

Dependent variable	(1) Absence rate	(2) Unpaid overtime	(3) Absence rate	(4) Unpaid overtime
Personal service occupations	0.022***	-1.815***	0.022***	-1.815***
	(0.003)	(0.061)	(0.003)	(0.061)
Sales and customer service	0.009***	-1.903***	0.009***	-1.907***
	(0.003)	(0.061)	(0.003)	(0.061)
Process, plant and machine	0.014***	-2.071***	0.014***	-2.072***
	(0.003)	(0.063)	(0.003)	(0.063)
Elementary occupations	0.012***	-1.941***	0.012***	-1.945***
	(0.003)	(0.058)	(0.003)	(0.058)
Age	0.001**	0.023***	0.001**	0.024***
	(0.000)	(0.006)	(0.000)	(0.006)
Age squared	-0.000	-0.000***	-0.000	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Male (Female exc.)	-0.003	-0.401***		
	(0.002)	(0.047)		
Qualification (degrees exc.)				
Higher education	-0.004*	-0.633***	-0.004*	-0.633***
	(0.002)	(0.060)	(0.002)	(0.060)
A level or equivalent	-0.003	-0.963***	-0.003	-0.961***
	(0.002)	(0.048)	(0.002)	(0.048)
GCSE or equivalent	-0.001	-1.001***	-0.001	-1.003***
	(0.002)	(0.047)	(0.002)	(0.047)
Other qualifications	-0.002	-1.116***	-0.002	-1.120***
	(0.002)	(0.052)	(0.002)	(0.052)
No qualification	0.001	-1.217***	0.001	-1.223***
	(0.003)	(0.050)	(0.003)	(0.050)
Marital status (Single exc.)				
Married	0.001	0.035	0.001	0.028
	(0.002)	(0.038)	(0.002)	(0.038)
Separated	0.000	-0.042	0.000	-0.029
	(0.004)	(0.072)	(0.004)	(0.072)
Divorced	0.007**	0.098*	0.007**	0.098*
	(0.003)	(0.056)	(0.003)	(0.056)
Widowed	0.013*	-0.020	0.012*	-0.035
	(0.007)	(0.095)	(0.007)	(0.096)
Dep. children in family	-0.000	-0.117***		
	(0.001)	(0.031)		
Paid overtime work hours		-0.068***		-0.068***
		(0.003)		(0.003)
Sex & dependent children (Male & no dep. child exc.)				
Male & dep. child			0.003	-0.075
			(0.003)	(0.068)
Female & no dep. child			0.004*	0.410***
			(0.003)	(0.055)
Female & dep. child			0.002	0.285***
			(0.003)	(0.062)
Constant	0.005	2.348***	0.001	1.944***
	(0.008)	(0.137)	(0.008)	(0.136)
Region	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
Adj. R-squared	0.005	0.206	0.005	0.206
No. of obs.	78029	78302	78029	78302

Note: ***, **,* refer to the significance levels at 1, 5 and 10% respectively. Robust standard errors are given in parenthesis. The estimated coefficients for the regional and year dummies are missed from the table.

In column (3) of Table 4.6, the regression results for the absenteeism rate with the interaction terms between commuting time and 'sex – dependent children' dummies is presented. The estimated coefficient is the biggest for women with dependent children among the four groups while in the case of unpaid overtime work, the estimated coefficient is the lowest for women. Both of the results imply that childcare certainly imposes time-constraints on female workers so that they decrease their effort to a greater degree than any other group for a given increase in commuting time. Another interesting finding is that the presence of dependent children works in the opposite direction for male workers. Those with some dependent children have negative but insignificant estimates with both of the dependent variables, meaning that their effort levels do not really respond to the length of commuting time. It seems possible that male workers with dependent children are likely to be the main bread winners within their families and therefore they do not really adjust their effort levels according to how far they have to travel to work.

4.4.4. Reverse causality

Though it turns out that there is significant correlation between worker effort and commuting time and the impact of commuting time on measures of work effort are as predicted by the neoclassical model, one cannot be certain that the correlation is definitely causal and that the direction of causation runs only from commuting time to work effort. Theoretically, worker effort can affect commuting time. Less work-oriented workers might choose longer commutes, for example, because they want to live in suburban areas as they care more about family, child care and quality of life than work. Therefore, it is possible that the observed

correlation between them is causal in both directions.

However, at least, it seems plausible to argue which of the two directions of causation contribute more to generating the observed correlation. Arguably, although less work-oriented individuals may want to live a long way from central business districts, they might also be expected to prioritise finding jobs near their homes since they are less work-motivated. In contrast, it is much more plausible that long commuting time causes workers to feel tired and hence negatively affect their willingness to exert effort at work. In the end, however, to infer the causation of interest, more sophisticated econometric approaches, such as IV estimation, should be tried. In this instance, a valid instrument should be correlated with commuting time but not with omitted variables which affect work effort. In theory, natural topology, such as rivers and mountains in the local areas, can be valid instruments as they can affect commuting time but they may not be correlated with unobservable characteristics that affect worker effort. Unfortunately, this identification strategy is not feasible with the data at hand as it does not provide information on detailed commuting routes of workers and therefore more suitable datasets need to be found for this strategy. A completely different but more feasible approach to identification of the causation may be to exploit a quasi-natural experiment which brings an exogenous change in commuting time of workers, such as new road construction, the London Underground upgrade, etc. Then, one can judge whether the exogenous change in commuting time causes a change in work effort. This strategy should be tried in the future with appropriate data.

4.5. Conclusion

This chapter explores the relationship between commuting time and workers' effort. Absenteeism and unpaid overtime work hours are identified as proxy measures of workers' effort through the literature review. The adoption of two measures reinforces the reliability of the empirical results. The dataset used for the empirical analysis is the QLFS which provides all the variables needed. When the relevant personal and demographical characteristics and job characteristics are controlled for, commuting time shows a positive correlation with absenteeism and a negative correlation with unpaid overtime hours as predicted through the neoclassical labour supply model.

The robustness of the results is checked through the regression for sub-sample groups using the interaction terms. The model predicts that a reduction in part-time worker effort levels is more pronounced for a given increase in commuting time as the increase lowers their effective wage to a greater extent than full-time workers. The empirical results turn out to be consistent with this prediction. Furthermore, the relationship for male and female workers is examined separately. It is reasonable to assume that female workers are more time-constrained as a result of family and childcare responsibilities and therefore their effort levels are expected to respond to the increase in commuting time more sensitively and this is empirically demonstrated to be true. Working women are more likely to be absent from work and do less unpaid overtime work than otherwise comparable men as their commuting time increases. This tendency is particularly stronger for women with dependent children than those without any.

The findings in this chapter have important implications for policy-makers, employers and workers. Firstly, when policy-makers weigh up the costs and benefits of building public transportation infrastructure (road, rail, etc.), they need to consider the benefits from the perspective of increased effort and productivity of workers. Secondly, employers could find it worth taking action to shorten employees' commuting time or to improve commuting conditions. This could be done, for example, by providing free shuttle buses or by subsidising public transportation costs or fuel for car-users as the company can benefit from increased worker effort. Thirdly, for workers themselves, as commuting time affects their work incentive and performance negatively, it can also harm promotional opportunities and prospects of pay rises and raise the likelihood of dismissal. Therefore, it would be worth considering moving closer to the workplace if gains from the reduced commuting time are greater than the costs of relocation.

CHAPTER 5. CONCLUSION

This thesis presented three empirical essays on the interaction of labour and housing markets. These two markets are geographically overlapping with each other and indeed they are likely to be coincident for most households as people want work and their other daily activities to occur within a relatively limited area for the efficiency of time and resources. This, indeed, is the very basis of the monocentric model of urban land use which is the foundation of modern urban economics. As the spatial extent of the two markets is similar, many joint social and economic outcomes arise between them. Therefore, it is difficult to fully understand housing market phenomena without taking the labour market into consideration and vice versa. This thesis focuses on and confirms a few of the important ways these two markets interact.

5.1. Discussion of the findings

Chapter 2 examines the causal relationship between one's tenure choice and unemployment status: specifically the proposition that homeownership causally and positively increases the probability of a person being unemployed. It has been argued that homeowners are less likely to move to other regions in response to negative demand shocks to their local labour markets due to large transaction costs in housing markets and therefore they are more likely to be unemployed at any given point in time. This is an important question for policy-makers as homeownership subsidies might have to be reconsidered if it were to turn out that there was this kind of causal relationship between homeownership and unemployment. The empirical challenge for the study comes from the possible

endogeneity of homeownership, which, if present, would cause OLS or logit estimators to be biased. To deal with any endogeneity, this research employs two plausible instruments, namely the local homeownership rate and parental homeownership status.

The local homeownership rate is considered a comprehensive measure of how accessible homeownership is in the local area and available from UK panel data (BHPS). Using the panel structure of the dataset, a FE IV model is estimated and no causal relationship between homeownership and unemployment is found. In the related literature, parental homeownership status is found to affect children's homeownership status positively through the inter-generational transfer of wealth and knowledge of the home-buying process. The British Cohort Study 1970 provides information on parental homeownership status when the cohort members were very young. When the cohort member's homeownership status is instrumented by parents' homeownership status, no causal effect of homeownership on unemployment is found.

Chapter 3 is the first empirical study to shed some light on the relationship between movers' tenure decisions and the information available on their destination housing markets. If it is allowed that the longer the distance moved, the more difficult it becomes for households to collect information on their destination housing markets, distance moved is a useful proxy for information on those markets. So, this leads to the testable proposition that mover households will be more likely to rent the further they have moved. This chapter tests and confirms three hypotheses each of which is related to the notion that local housing

market information is important for movers' tenure decisions. Firstly, the amount of information on destination housing markets held by movers decreases with the distance moved when it is proxied for by movers' awareness of problems (e.g. crime, vandalism) in the new neighbourhoods. Secondly, movers tend to rent as the amount of information on the destination housing market decreases and that this is represented by the distance moved. Lastly, the longer the distance moved, the more quickly the households move again. This finding is consistent with the notion that the speed of the locational adjustment within the current housing market is expected to increase with the distance moved as longer-distance movers are less likely to be satisfied with the initial accommodation and neighbourhoods due to the lack of information on them prior to moves. Persuasive evidence is found in support of all three of these hypotheses.

Chapter 4 demonstrates the negative impact of commuting time on worker effort. This issue has important implications for the various parties affected by worker effort. Despite the potentially important implications, the topic has been largely unexplored until a recent study with German data which confirmed the effect of commuting time on labour productivity. Though absenteeism rates and unpaid overtime hours have little in common except that they are closely related with worker effort, when they are adopted as a proxy measure for worker effort, they lead to the same conclusion that commuting time affects worker effort negatively. Even when the effect of commuting time on work effort is compared between sub-groups of the regression sample (part-time and full-time, male and female workers), the results are consistent with the predictions made based on the neoclassical labour supply theory.

5.2. Contributions, limitations and future studies

The most notable contribution made by this thesis is that it investigates relatively unexplored topics on the interaction between housing and labour markets – the two most spatial markets. Chapter 3 is the first empirical test of whether the distance moved affects movers' tenure decisions. Though this topic can have significant implications for the rental market and labour mobility, there has been little attempt to evaluate the influence of local information on tenure decisions. This is probably because of the difficulty of framing the question in a way that would allow it to be addressed with available data. A significant contribution of this thesis is that it overcomes the difficulty of measuring the amount of local housing market information held by households prior to their moves by looking at the distance they moved. Chapter 4 is the first UK study of the impact of commuting time on worker effort. There have been many studies which measure the negative impact of commuting on worker health or well-being but none on worker effort or productivity directly. Considering the significant implications of this issue to the various parties concerned with worker effort in the workplace, it is surprising that it has not previously been the subject of research.

Unlike the other two chapters, Chapter 2 focuses on a relatively well-explored topic and its contribution to the literature is mainly methodological. To identify the causal effect of an individual's homeownership on unemployment status, it employs an IV approach with the two different types of instruments (local homeownership rate and parental homeownership status) which have not been applied in previous British studies. Especially, it complements existing studies in the sense that it is the first in the UK to combine the FE model and IV approach in

order to more rigorously assess the causal effect of homeownership on unemployment.

Compared to Chapter 2 which adopts an IV approach to deal with the endogeneity of the dependent variable, Chapter 4 pays relatively less attention to the possibility of reverse causality and how to deal with it. It is concluded that commuting time has a negative impact on worker effort but theoretically the latter could also determine the former. For instance, less work-oriented workers may choose to live far away from their jobs to enjoy what they judge to be a better living environment. Though the reverse causality is not expected to be strong enough to invalidate the conclusion of Chapter 4, related further work might try to take a more sophisticated econometric approach to identify the direction of causality more precisely. One way to identify causation might be to exploit quasi-natural experiments such as the new construction of roads or rail links or upgrades such as to London Underground, which cause an exogenous change in commuting time for workers in the affected areas.

The obvious limitation of Chapter 2 is that it alone cannot tell *why* there is no relationship between unemployment and homeownership. However, the result of 'no causal relationship' does not necessarily mean that nothing goes on between homeownership and unemployment. Homeowners may behave differently from renters to lower their chance of becoming unemployed or adopt different search strategies. For example, Munch et al. (2006) suggest that homeowners are more likely than renters to accept local jobs with low wage rates rather than stay unemployed. As there now seems to be a general consensus that homeownership

status does not cause unemployment, it is more important to understand why homeowners are not necessarily more likely to be unemployed than renters despite their obviously lower residential mobility. To date, it is this more general issue which has been relatively overlooked. In particular, it should be carefully examined whether homeowners are disadvantaged in terms of wage, commuting time, unpaid overtime work, job-skill mismatch and so on to retain their jobs. If homeowners maintain their employment only at the cost of decent working conditions, the subsidy schemes for homeownership may still need to be reconsidered even though it does not have a negative impact on employment status.

The findings from a recent study raises a more fundamental question regarding Chapter 2: Are homeowners really less mobile than renters even when they want to move for employment opportunities? Hilber and Lyytikäinen (2012) show that stamp duty in the UK tends to hamper homeowner mobility in general but job-related movers are an exception. This may be because the gains from job-related moves (e.g. higher wages and better chances of employment) exceed the moving costs. In contrast, Munch et al. (2006) show that homeownership lowers the likelihood of moving for job reasons. Therefore, it is necessary to examine carefully whether homeowners are less mobile than renters when they need to move for jobs and whether the degree of the homeowner's job-related mobility varies by country or characteristics.

One more issue worth further consideration in relation to Chapter 2 is why there is a gap between macro-level and micro-level studies of the relationship between

unemployment and homeownership. The results from the micro-studies exclude the possibility that the relationship between individuals' homeownership and unemployment status is an underlying mechanism for the positive correlation between homeownership and unemployment rates at the macro-level. Therefore, it seems necessary to investigate what causes a positive correlation between those two rates and whether it has any important implications for housing and labour market policies.

The implication of Chapter 3 for the labour market is that the availability of easily accessible accommodation in the housing/labour markets to which job-related movers want to move would be one of the most important factors to consider for the moving and location decisions for them. Based on this implication, a suggestion for further study related to Chapter 3 is to look at whether the existence of a larger private rental sector in a housing market would increase interregional mobility and attract more people to the local labour market and hence promote labour mobility at the national level. If it is supported empirically, some policy measures to encourage the growth of a more flexible and affordable private rental sector to improve the efficiency of national labour markets might be worth considering.

REFERENCE

Aaronson, D. (2000) A note on the benefits of homeownership, *Journal of Urban Economics,* vol. 47, pp.356-369.

Anger, S. (2005) Working time as an investment? - The effect of unpaid overtime on wages, promotions and layoffs, DIW, Discussion Paper no. 535, Berlin.

Barmby, T. (2002) Worker absenteeism: A discrete hazard model with bivariate heterogeneity, *Labour Economics*, vol. 9, pp. 469-476.

Barmby, T., A. Marco, G. Ercolani and J. Treble (2002) Sickness absence: An international comparison, *Economic Journal*, vol. 112, F315-331.

Barmby, T., C. Orme and T. John (1991) Worker absenteeism: An analysis using microdata, *Economic Journal*, vol. 101, pp. 214-229.

Barmby, T., C. Orme and T. John (1995) Worker absence histories: A panel data study, *Labour Economics*, vol. 2, pp. 53-65.

Barrios Garcia, J. and J. Rodriguez Hernandez (2004) User cost changes, unemployment and home-ownership: Evidence from Spain, *Urban Studies*, vol. 41, pp. 563-578.

Battu, H., A. Ma and E. Phimister (2008) Housing tenure, job mobility and unemployment in the UK, *Economic Journal*, vol. 18, pp. 1-18.

Boehm, T. and A. Schlottman (1999) Does homeownership by parents have an economic impact on their children?, *Journal of Housing Economics*, vol. 8, pp.217-232.

Booth, A., M. Francesconi and J. Frank (2002) Temporary jobs: Stepping stones or dead ends?, *Economic Journal*, vol. 112, F189-F213.

Borjas, G. (2005) *Labor Economics*, New York: McGraw-Hill.

Charles, K. and E. Hurst (2002) The transition to home ownership and the black-white wealth gap, *Review of Economics and Statistics*, vol. 84, pp. 281-297.

Clark, A. and Y. Huang (2004) Liking migration and mobility: Individual and contextual effects in housing markets in the UK, *Regional Studies*, vol. 38, pp. 617-628.

Coulson, N. and L. Fisher (2002) Labor market outcomes of homeownership, *Housing Studies*, vol. 17, pp. 35-50.

Coulson, N. and L. Fisher (2009) Housing tenure and labor market impacts: The search goes on, *Journal of Urban Economics*, vol. 65, pp. 252-264.

Courchane, M., B. Surette and P. Zorn (2004) Subprime borrowers: Mortgage transitions and outcomes, *Journal of Real Estate Finance and Economics*, vol. 29, pp. 365-392.

DaVanzo, J. (1983) Repeat migration in the United States: Who moves back and who moves on?, *Review of Economics and Statistics*, vol. 65, pp. 552-559.

Diaz-Serrano, L. (2005) Labor income uncertainty, skewness and homeownership: A panel data study for Germany and Spain, *Journal of Urban Economics*, 2005, vol. 58, pp.156-176.

Deutsch, E. (1997) Indicators of housing finance intergenerational wealth transfers, *Real Estate Economics*, vol. 25, pp.129-172

Dietz, R., and D. Haurin (2003) The social and private micro-level consequences of homeownership, *Journal of Urban Economics*, vol. 54, pp. 401-450.

Dujardin, C. and F. Goffette-Nagot (2008) Does public housing occupancy increase unemployment?, Working Paper, GATE Groupe d'Analyse et de Théorie Économique.

Engellandt, A. and R. Riphahn (2005) Temporary contracts and employee effort, *Labour Economics*, vol. 12, pp. 281-299.

Engelhardt, G. and C. Mayer (1994) Gifts for home purchase and housing market behavior, *New England Economic Review*, vol. 94, pp. 47-58.

Evans, G. and R. Wener (2006) Rail commuting duration and passenger stress, *Health Psychology*, vol. 25, pp. 408-412.

Farley, R. (1996) Racial differences in the search for housing: Do Whites and Blacks use the same techniques to find housing?, *Housing Policy Debate*, vol. 7, pp. 367-385.

Filer, R., D. Hamermesh and A. Rees (1996) *The economics of work and pay*, New York: HarperCollins College Publishers.

Flatau, P., M. Forbes, P. Hendershott and G. Wood (2003) Homeownership and unemployment: The roles of leverage and public housing, NBER working paper 10021.

Frees, E. (2004) *Longitudinal and panel data: Analysis and applications in the social sciences*, CUP: New York.

Fu, Y. (1991) A model of housing tenure choice: comment, *American Economic Review*, vol. 81, pp. 381-383.

Fu, Y. (1995) Uncertainty, liquidity, and housing choices, *Regional Science and Urban Economics*, vol. 25, pp. 223-236.

Garcia, R. (1987) Sick time usage by management and professional employees in the public sector, *Review of Public Personnel Administration*, vol.7, pp. 45-49.

Gordon, P., H. Richardson and M. Jun (1991) The commuting paradox: Evidence

from the top twenty, *Journal of the American Planning Association*, vol. 57, pp. 416-420.

Green, R. and M. White (1997) Measuring the benefits of homeowning: Effects on children, *Journal of Urban Economics*, vol. 41, pp. 441-461.

Green, R. and P. Hendershott (2001) Homeownership and unemployment in the US, *Urban Studies*, vol. 38, pp. 1509-1520.

Ha, S. (2005) *Do worker's effort decrease with commuting time?*, Unpublished master dissertation, London School of Economics, London, UK.

Hamnett, C., M. Harmer and P. William (1991) *Save as houses: Housing inheritance in Britain*, Philip Chapman Publishing: London.

Haurin, D. (1991) Income variability, homeownership, and housing demand, *Journal of Housing Economics*, vol. 1, pp. 60-74.

Haurin, D. and H. Gill (1987) Effects of income variability on the demand for owner-occupied housing, *Journal of Urban Economics*, vol. 22, pp. 136-150.

Haurin, D. and H. Morrow-Jones (2006) The impact of real estate market knowledge on tenure choice: A comparison of black and white households, *Housing Policy Debate*, vol. 17, pp. 625-653.

Haurin, D., T. Parcel and R. Haurin (2001) The impact of homeownership on child outcomes, Working Paper, Joint Center for Housing Studies, Harvard University.

Heckman, J. and B. Singer (1984) A method for minimizing the impact of distributional assumptions in econometric models for duration data, *Econometrica*, vol. 52, pp.271-320.

Hendershott, P. and M. White (2000) The rise and fall of housing's favored investment status, *Journal of Housing Research*, vol. 11, pp. 257-275.

Henretta, J. (1984) Race differences in middle-class lifestyle: The role of homeownership, *Housing Economics*, vol.1, pp. 60-74.

Hilber, C. (2005) Neighborhood externality risk and the homeownership status of properties, *Journal of Urban Economics*, vol. 57, pp. 213-241.

Hilber, C. and Y. Liu (2008) Explaining the black-white homeownership gap: The role of own wealth, parental externalities and locational preferences, *Journal of Housing Economics*, vol. 17, pp. 152-174.

Hilber, C. and T. Lyytikäinen (2012) Stamp duty and household mobility: regression discontinuity evidence from the UK, LSE, London, UK.

Hirad, A. and P. Zorn (2002) Pre-purchase homeownership counselling: A little knowledge is a good thing. In *Low-Income Homeownership*, ed. Nicolas P. Restinas and E.S. Belsky, pp. 146-175, Brookings Institution Press: Washington, DC.

Hoehner, C., C. Barlow, P. Allen and M. Schootman (2012) Commuting distance, cardiorespiratory fitness and metabolic risk, *American Journal of Preventive Medicine*, vol. 42, pp. 571-578.

Hughes, G. and B. McCormick (2000) *Housing policy and labour market performance*, Department of the Environment, Transport and the Regions: London.

Johannson, P. and M. Palme (1996) Do economic incentives affect work absence? Empirical evidence using Swedish micro data, *Journal of Public Economics*, vol. 59, pp. 195-218.

Johannson, P. and M. Palme (2002) Assessing the effect of public policy on worker absenteeism, *Journal of Human Resources*, vol. 37, pp. 381- 409.

Johns, G. (1978) Attitudinal and non-attitudinal predictors of two forms of absence from work, *Organizational Behavior and Human Performance*, vol. 22, pp. 431-444.

Kain, J. (1968) Housing segregation, Negro employment and metropolitan decentralization, *Quarterly Journal of Economics*, vol. 82, pp. 175-197.

Koslowsky, M., A. Kluger, M. Reich (1995) *Commuting stress: Causes, effects and methods of coping,* New York: Plenum Press.

Lazear, E. and S. Rosen (1990) Male-female wage differentials in job ladders, *Journal of Labor Economics*, vol. 8, S106-123.

Lee, B. and J. McDonald (2003) Determinants of commuting time and distance for Seoul residents: The impact of family status on the commuting of women, *Urban Studies*, vol. 40, pp. 1283-1302.

Manning, A. (2003) The real thin theory: Monopsony in modern labour markets, *Labour Economics*, vol. 10, pp. 105-133.

Mastekaasa, A. and K. Olsen (1998) Gender, Absenteeism and job characteristics: A fixed effects approach, *Work and Occupations*, vol. 25, pp. 195-228.

Mulder, C. and J. Smits (1999) First-time home-ownership of couples - The effect of intergenerational transmission, *European Sociological Review*, vol. 15, pp.323-337.

Munch, J., M. Rosholm and M. Svarer (2006) Are homeowners really more unemployed?, *Economic Journal*, vol. 116, pp. 991-1013.

Munch, J., M. Rosholm and M. Svarer (2008) Home ownership, job duration, and wages, *Journal of Urban Economics*, vol. 63, pp. 130-145.

Munro, M. (1989) Housing wealth and inheritance, *Journal of Sociology and Politics*, vol. 17, pp.417-436.

Nickell, S. (1998) Unemployment: Questions and some answers, *Economic Journal*, vol. 108, pp. 802-816.

Nivalainen S. (2004) Determinants of family migration: short moves vs. long moves, Journal of Population Economics, vol. 17, pp. 157-175.

Oswald, A. (1996) A conjecture of the explanation for high unemployment in the industrialized nations: Part I, Paper No. 475, Warwick University Economic Research.

Oswald, A. (1999) The housing market and Europe's unemployment: A non-technical paper, Warwick University.

Pannenberg, M. (2002) Long-term effects of unpaid overtime: Evidence for West Germany, IZA, Discussion Paper no. 614., Bonn.

RAC Foundation (2007) *The UK commute: Healthy or hazardous? An analysis of current commuting trends*, London, UK.

Pehkonen, J. (1999) Unemployment and home-ownerhip, *Applied Economic Letters*, vol. 6, pp.263-265.

Robst, J., R. Deitz and K. McGoldrick (1999) Income variability, uncertainty and housing tenure choice, *Regional Science and Urban Economics*, vol. 29, pp. 219-229.

Rocheteau, G. (2000) Equilibrium unemployment and wage formation with matching frictions and worker moral hazard, *Labour Economics*, vol. 8, pp. 75-102.

Rosen, H., K. Rosen and D. Holtz-Eakin (1984) Housing tenure, uncertainty, and taxation, *Review of Economics and Statistics*, vol. 66, pp. 405-416.

Rouwendal, J. and E. Meijer (2001) Preferences for housing, jobs, and commuting: A mixed logit analysis, *Journal of Regional Science*, vol. 41, pp. 475-505.

Rouwendal, J and P. Nijkamp (2004) Living in two worlds: A review of home-to-work decisions, *Growth and Change*, vol. 35, pp.287-303.

Rouwendal, J. and P. Nijkamp (2010) Homeownership and labour market behaviour: Interpreting the evidence, *Environment and Planning A*, vol. 42, pp. 419-433.

Schwanen, T. and D. Martin (2002) Travel-time ratios for visits to the workplace: the relationship between commuting time and work duration, *Transportation Research Part A*, vol. 36, pp. 573-592.

Shapiro, C. and J. Stiglitz (1984) Equilibrium unemployment as a discipline device, *American Economic Review*, vol. 74, pp. 433-444.

Tatch, J. (2007) Affordability - are parents helping?, *CML Housing Finance*, Issue 03.

Turner, T. (2003) Does investment risk affect the housing decisions of families?, *Economic Inquiry*, vol. 41, pp. 675-691.

Turner, T. and D. Seo (2007) Investment risk and the transition into homeownership, *journal of Regional Science,* vol. 47, pp. 229-253.

Van Leuvensteijn, M. and P. Koning (2004) The effect of home-ownership on labor mobility in the Netherlands, *Journal of Urban Economics,* vol. 55, pp. 580-596.

Van Ommeren, J., P. Rietveld and P. Nijkamp (1997) Commuting: In search of jobs and residences, *Journal of Urban Economics*, vol. 42, pp. 402-421.

Van Ommeren, J. and E. Gutiérrez-i-Puigarnau (2011) Are workers with a long commute less productive? An empirical analysis of absenteeism, *Regional Science*

and Urban Economics, vol. 41, pp. 1-8.

Wener, R., G. Evans and P. Boately (2006) Commuting stress: Psychophysiological effects of a trip and spillover in the workplace, *Transportation Research Record: Journal of the Transportation Research Board*, no. 1924, pp. 112-117.

Yezer, A. and L. Thurston (1976) Migration patterns and income change: Implications for the human capital approach to migration, *Southern Economic Journal*, vol. 42, pp. 693-702.

Zenou, Y. (2002) How do firms redline workers?, *Journal of Urban Economics*, vol. 52, pp. 391-408.