ABSTRACT

Title of dissertation: STUDYING THE CAUSES AND CONSEQUENCES OF INTERNAL LABOR MIGRATION USING SURVEY AND ADMINISTRATIVE DATA SOURCES

Christopher F. Goetz, Doctor of Philosophy, 2016

Dissertation directed by: Professor John C. Haltiwanger
Department of Economics

This dissertation comprises three chapters. The first chapter motivates the use of a novel data set combining survey and administrative sources for the study of internal labor migration. By following a sample of individuals from the American Community Survey (ACS) across their employment outcomes over time according to the Longitudinal Employer-Household Dynamics (LEHD) database, I construct a measure of geographic labor mobility that allows me to exploit information about individuals prior to their move. This enables me to explore aspects of the migration decision, such as homeownership and employment status, in ways that have not previously been possible. In the second chapter, I use this data set to test the theory that falling home prices affect a worker's propensity to take a job in a different metropolitan area from where he is currently located. Employing a within-CBSA and time estimation that compares homeowners to renters in their propensities to relocate for jobs, I find that homeowners who have experienced declines in the nominal value of their homes are approximately 12% less likely than average to take a

new job in a location outside of the metropolitan area where they currently reside. This evidence is consistent with the hypothesis that housing lock-in has contributed to the decline in labor mobility of homeowners during the recent housing bust. The third chapter focuses on a sample of unemployed workers in the same data set, in order to compare the unemployment durations of those who find subsequent employment by relocating to a new metropolitan area, versus those who find employment in their original location. Using an instrumental variables strategy to address the endogeneity of the migration decision, I find that out-migrating for a new job significantly reduces the time to re-employment. These results stand in contrast to OLS estimates, which suggest that those who move have longer unemployment durations. This implies that those who migrate for jobs in the data may be particularly disadvantaged in their ability to find employment, and thus have strong short-term incentives to relocate.

STUDYING THE CAUSES AND CONSEQUENCES OF INTERNAL LABOR MIGRATION USING SURVEY AND ADMINISTRATIVE DATA SOURCES

by

Christopher F. Goetz

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2016

Advisory Committee:
Professor John Haltiwanger, Chair
Professor Katharine Abraham
Professor Judith Hellerstein
Professor John Shea
Professor Phillip Swagel

Preface

This document reports the results of research and analysis undertaken while the author was employed by the U.S. Census Bureau. All views expressed herein are those of the author and do not necessarily represent the views of the U.S. Census Bureau or the Department of Commerce. Most of the research contained in this dissertation makes use of confidential data, which is available to external researchers through the Federal Research Data Center network (http://www.census.gov/fsrdc). All results have been reviewed to ensure that no confidential information has been disclosed. For further information regarding the LEHD data, please contact the Center for Economic Studies, U.S. Census Bureau, 4600 Silver Hill Road, Suitland Maryland, 20746, USA (CES.Local.Employment.Dynamics@census.gov).

Dedication

In loving memory of Joëlle B. Goetz.

Acknowledgments

I would like to express my sincere gratitude to the members of my dissertation committee for their input and guidance, and especially to my advisor John Haltiwanger for leading me through the process. This research also benefited greatly from help from my colleagues at the Center for Economic Studies at the U.S. Census Bureau.

Finally, I would like to thank my family, particularly my wife Mariel, for their continued love and support over the years.

Table of Contents

List of Tables

List of Figures

Chapter 1: The Potential for Using Combined Survey and Administrative Data Sources to Study Internal Labor Migration

1.1 Introduction

The idea of America as a land of opportunity has long been associated with the vision of people moving freely across the country in order to seek a better fortune. While international comparisons are difficult, the U.S. population has generally been considered to have one of the highest mobility propensities in the world, which has been viewed as a hallmark strength of the labor market (Long, 1991). This depiction belies the fact, however, that migration has been experiencing a secular decline for the past 35 years or so, as well as a precipitate decline during the Great Recession, leading to much research into the causes (Molloy, Smith, and Wozniak, 2011). Two leading culprits are the labor market, which has also been demonstrating a secular decline in other types of worker transitions, and the housing market, which suffered a historic collapse following the price bubble of the mid-2000s. The inability of the unemployed to relocate has been a particular source of concern, since this group lacks the resources to withstand falling property values, and may also suffer disproportionately from being unable to move away from distressed areas with high jobless rates.

Research on these topics has typically been conducted using surveys such as the Decennial Census and the Current Population Survey (CPS), or administrative data from income tax records provided by the Internal Revenue Service (IRS). In the studies that follow, I create a novel data set that combines survey data from the American Community Survey (ACS), with administrative data on employment outcomes from the Longitudinal Employment-Household Dynamics (LEHD) data base, both created by the U.S. Census Bureau. With this novel data set I am able to address a couple of questions in the migration literature whose study has suffered from certain data limitations. First, I test the much discussed theory that the crash in the housing market contributed to the decline in migration, by leaving homeowners with insufficient home equity to sell their houses and move. Secondly, I address whether the recent decline in migration should truly be a concern for the efficiency of the labor market, by testing whether changing labor markets has a positive impact on the job prospects of the unemployed, a group for whom migration is posited to be an important means by which to escape joblessness.

Combining survey data with administrative records in this way allows me to leverage the strengths of each data source, and in turn compensate for the shortcomings of the other. Survey data, with a rich set of information about individuals, provides the variables necessary to address a myriad of economic topics of interest. However, due to the financial and logistical constraints of conducting surveys, these data sets seldomly record the same entities longitudinally, and thus deny the researcher the ability study the future outcomes in which they may be interested. The longitudinal surveys that exist, such as the Survey of Income and Program Participation (SIPP) and National Longitudinal Surveys (NLS), have relatively small sample sizes that often do not afford the econometrician sufficient statistical precision to test their hypotheses. Administrative records, on the other hand, are more narrow in focus and lack the detailed information about the economic agents. However, because these data are recorded routinely and as a matter of course for the administration of large programs or efforts, they typically have a more comprehensive scope spanning multiple years, often observing the same individuals on a repeated basis.

This type of data proves particularly beneficial in researching migration, an economic outcome that lends itself to longitudinal study. Existing surveys that do include a longitudinal component, such as the CPS, do not follow an individual after they leave their residence, thus it is not possible to determine where the individual moved to. The CPS, ACS, and Decennial Census all include retrospective measures of migration, asking a respondent whether they moved in the past year, or 5 years. While this provides a measure of migration, it does not provide any information about the respondent's situation during the time preceding and during their move. While obviously some characteristics of an individual are consistent over time, like their sex, race, and date of birth, many others could well have changed. For example, when it comes to homeownership, it is not an obvious assumption that someone who is currently observed owning or renting did the same prior to being surveyed. Likewise, it is not possible to know from these surveys what an individual's employment status was in their prior location. The LEHD data, on the other hand, is an administrative database that allows researchers to track individuals at different jobs in different locations over time, but lacks much other information about the worker and their household. Clearly, a great potential exists for bringing these two types of of data together, in order to combine the unique details of a survey with a means of observing subsequent labor migration via the administrative records.

In this first introductory chapter, I discuss the data sets that typically have been used to study migration, and motivate how my newly created data set can address some shortcomings in the previous literature. I also present some basic measures of migration from this data set, in order to compare them to the statistics derived from existing sources. Next, the main analytical chapter will describe the construction of the new data set in greater detail, and show how it can be employed in testing the hypothesis that declining house prices and negative equity lead to lower mobility propensities of homeowners. Results from the analysis reveal that workers are significantly less likely to be observed relocating if they have experienced a nominal price decline since purchasing their home, although the impact on the unemployed appears to be somewhat mitigated. In the final chapter, I address the question of whether or not we should truly be concerned about the recent decline in migration, by testing whether the unemployed who migrate for their next job have shorter unemployment spells than they would have by remaining in their original location. In contrast to previous findings, results which control for selection into migration show that workers who move for new employment are more likely to be employed in a shorter timespan than had they stayed in the same location, despite the fact that their unemployment spells are somewhat longer on average. This may imply that those who decide to move are the most disadvantaged workers, and thus may stand the most to gain from relocating.

The rest of the current chapter proceeds as follows. Section 2 discusses the data sets that are typically employed in the literature on internal migration, and introduces the idea behind the construction of my combined survey/administrative database. Section 3 discusses the broad topics in migration that the data have been applied to, and discusses how my new data set can help address some of the shortcomings in the exisiting literature. Section 4 discusses how migration is typically measured, and provides comparisons between statistics derived from my new data set with those from other sources. Section 5 concludes.

1.2 Data Sets Used to Study Migration

Researchers have principally used three data sets to obtain precise, nationally-representative measures of migration in the United States: the Decennial Census, and more recently its annual supplement, the ACS; the Annual Social and Economic Supplement (ASEC) of the CPS; and the IRS origin-destination data.

The decennial census has provided information on migration status since 1940, asking respondents whether they lived in the same residence 5 years prior, and if not, what state they lived in at that time. This question began asking the county of previous residence starting in 1980. The ACS, which began in a beta version in 2000 and full scale starting in 2005, provides the same information as the census but on an annual frequency. As of 2010, the ACS migration question has supplanted the one on the decennial census, as the ACS was designed to eliminate the need for the decennial "long-form". The census and ACS also provide a way to measure long-term migration, as well as return migration, by providing the individual's state of birth. These data provide a rich details about an individual's demographics, employment status, income, and household conditions, but do not have a longitudinal component aside from the retrospective migration question. Decennial censuses can be linked using the confidential microdata, but the ten-year gap between observations is problematic for many research questions. As the ACS is a cross-sectional sampling of approximately 5 million respondents, repeated sampling of the same individual is rare.

The CPS microdata has contained a migration variable in its ASEC supplement since 1965, although migration rates have been published using the CPS going back to 1948. The CPS contains essentially the same information as the ACS, asking respondents where they lived in the previous year, down to the county level. The CPS is the most detailed of the Census-based surveys with a host of information about employment, income, health insurance and employer benefits. However, the CPS has a relatively small sample size, interviewing only about 100,000 households annually compared to 3 million in the ACS, rendering it less suitable for more granular measurements and data-intensive analyses. Also, while respondents in the CPS are interviewed multiple times, those who leave their residence during the interview window are not followed to their destination, thus preventing detection of subsequent migration. Nevertheless, some studies have used the fact of premature departure from the sample as a measure of residential mobility.

The IRS origin-destination migration statistics have been released since 1975, providing measures of county-to-county and state-to-state population flows. These data are derived from individual tax returns, which provide the address of primary residence for each household. Migration is defined as a year-to-year change in the location of primary residence, weighted by the number of people in each household according to the number of exemptions claimed on the return. The published statistics provide counts and rates of gross and net migration between all county pairs, making it the most comprehensive measure of household mobility. While it is not known whether tax filers are substantially different than non-tax filers in their migration propensities, it is estimated that approximately 90% of the population is covered. Being an administrative database, however, the IRS data does not include any of the detailed characteristics available in surveys. The confidential microdata that is available for purchase allows researchers to observe the income figures on the de-identified returns, but no further information about the individual or household is available.

A small number of longitudinal surveys exist that allow the user to observe individuals in their original location and conditions prior to migrating, as well as in their destination state. The National Longitudinal Survey of Youth 1979 (NLSY79) and Survey of Income and Program Participation (SIPP) are the two surveys that have been used the most for this purpose. While nationally representative, these surveys are limited due to their small sample sizes (under 50,000 survey units), which make detailed analyses and regression estimates less precise. Nevertheless, several studies have employed these data in the study of topics such as return migration, for example DaVanzo (1983), who found that over 25% of migrators are returning to places they had previously lived.

The new data set that I construct for my analyses combines many of the features of the survey and administrative sources described above, thus leveraging the strengths of each. Using the ACS as my sample population, I am able to take advantage of its large size and rich information about the individuals and households. By linking in information on subsequent employment from the LEHD jobs database, which covers over 95% of private employment in participating states, this allows near-comprehensive detection of future jobs across the nation, thus providing a measure of labor migration. This approach has several advantages over using either the ACS or the LEHD database alone. First, by recording things about the individual in their original location prior to moving, the data enable the researcher to more accurately account for the factors that enter into the migration decision. This is particularly important for the two topics that I address in this dissertation, namely homeownership and unemployment status, which clearly may vary across the origin and destination states. Secondly, it provides a direct measurement of labor mobility, in contrast to other sources that measure household mobility. After all, in terms of understanding the effect of migration on the labor market, we are most concerned about whether workers are able to efficiently match with firms located in other areas, regardless of whether a change in residences actually occurs. Also, it permits a more flexible definition of migration in terms of timing thresholds, especially due to the quarterly nature of the LEHD data. Instead of being bound by the ACS or IRS definition of annual migration, I can use higher-frequency measures as needed for the purposes of the given research question. Finally, the large size of the ACS sample provides high degree of precision for conducting many types of detailed analyses. This proves particularly useful for the analyses in the following chapters, as the ability to compare several individuals located in the same place and time allows me to control for many confounding labor market-related factors.

1.3 Migration Literature

The data described above began to be created at a time when theoretical models of migration were starting to take shape. Much of the early literature described the phenomenon as being driven by geographical differences in the supply and demand for labor. Seminal studies such as Lewis (1954) introduced a 2-sector model where equilibrium is restored by workers moving from the low-wage and less capital-intensive agriculture sector, to the high-wage and more capital-intensive manufacturing industry. These theories helped explain the movement from rural areas to cities that the U.S. had experienced earlier in the century, as well as spawned a new field of study in international development.

The micro-economic foundations of migration focus on the incentives of individual workers to choose the location that will maximize their utility. Todaro (1969) and Harris and Todaro (1970) described workers as seeking to maximize expected income, while taking into account the probability of unemployment. Sjaastad (1962) discussed the costs of migration more explicitly, stating that the increase in expected lifetime earnings must exceed both the monetary and non-monetary costs of migrating.

While this basic neo-classical framework has continued to hold for decades, in the latter part of the century the focus of the literature turned to identifying more specific determinants of migration, as summarized in Greenwood (1997). Many studies have explored the mobility of different demographic groups, showing evidence that young, higher skilled, and educated workers are more likely to move (Sjaastad, 1962; Borjas, 1991; Malamud and Wozniak, 2010). Household composition has also been considered, with married couples and families with children appearing less mobile in the data. The regional differences in amenity values has also been found to influence migration, including factors such as public spending and services (Tiebout, 1956), or natural amenities like climate and geography (Knapp and Graves, 1989; Mueser and Graves, 1995). Some studies have also focused on the cost side of the decision, including the increased burden from moving longer distances (Davies et al., 2001), and the risks associated with uncertainty (Schaeffer, 1988). Other financial factors, such as housing costs and mortgage rates, have also been found to be important (Cameron and Muellerbauer, 1998; Jackman and Savouri, 1992).

Research into the forces behind migration has become particularly important in light of the dramatic drop in mobility during the Great Recesssion, as well as the long-term decline over past decades, leading researchers to investigate possible explanations[1]. Compositional explanations have been sought, especially with the aging of the U.S. population and the fact that older workers tend to move less, but these types of differences across demographic groups are not of a sufficient magnitude to account for much of the overall decline (Molloy et al., 2011). The business cycle has been shown to be important with migration rates exhibiting pro-cyclical patterns, although this obviously does not explain the secular decline, nor

[1]This topic is reviewed extensively in Molloy et al., (2011).

does it go far in accounting for the precipitate decline during the Great Recession (Greenwood et al., 1997, Molloy et al., 2011).

One potential cause that has received a great amount of attention is the historic collapse of the housing market leading up to and during the Great Recession, the topic of my second chapter. The phenomenon of housing "lock-in" has been noted since Genesove and Mayer (1997; 2001) showed evidence that homeowners who had experienced a decline in their home equity were more reluctant in selling their homes. This gave support to the theory that owing more on one's mortgage than the home is worth, a condition called "negative equity" or being "underwater", makes homeowners less likely to relocate. The possible implications for long-distance migration led to an active literature in the topic, especially after the historic housing collapse of the mid-2000s. While some earlier studies provided evidence of lock-in hampering mobility, such as Chan (2001), Engelhardt (2003) and Ferreria, Gyourko, and Tracy (2008), more recent studies of the period surrounding the Great Recession and housing bust have tended to find no relationship (Dennet and Modestino, 2011; Molloy et al., 2011; Schulhofer-Wohl, 2012; Farber, 2012; Valletta, 2013).

This literature, however, suffers from a few limitations that my newly created data is well-suited to address. First of all, mobility in most of these studies is detected ex-post, and thus nothing is known about homeownership in the original location. While some studies make assumptions about prior homeownership, these analyses are flawed if former homeowners become renters in their new location in such a way that correlates with their prior negative equity status. Because my data set is constructed in such a way that allows us to observe an individual's migration outcome subsequent to being surveyed in their original location, I have direct information about their ex-ante homeownership and equity status. Secondly, all of the above studies depend on some kind of cross-geographic variation in housing prices, by city, state, or region. Thus they are unable to fully account for many of the differences between local labor markets, which may be correlated with local house prices and potentially bias the results. Unlike the data sets used in these prior studies, the data set I construct has both the geographic detail and the sample size required to exploit within-labor market variation in house prices and mobility, and thus compare owners and renters who are exposed to the same labor market at the same time. Finally, while the previous studies based on surveys or tax data measure residential mobility, the administrative jobs data that I use allow me to detect where their subsequent jobs are located, thus providing a direct measure of labor mobility. While residential and labor mobility are clearly correlated they are not necessarily the same thing, due to factors such as long-distance commuting and telework, and thus my data set allows us to observe the outcome most relevant to the labor market.

Implicit in all the preceding discussion of the literature is the assumption that migration is indeed a sign of a healthy and efficient labor market, and that anything that causes it to decline is therefore a harmful labor market friction. This has been a particular concern of economists in the wake of the Great Recession, who worry that the inability of the unemployed to pick up and move could prolong their unemployment spells and thus impede the economic recovery. However, the notion that workers who choose to move will be better off than had they remained is relatively underexplored in the literature. Kennan and Walker (2011) employ a structural estimation using the NLSY79 and find that locational decisions are consistent with the maximization of lifetime income, but they do not consider the unemployed in particular, or any shorter-term incentives that workers may have for moving. Studies focusing on the duration of unemployment spells of those who migrate compared to those who stay in place, have generally found zero or negative effects of migrating (Shumway, 2000; Pekkala and Tervo 2002). In the final chapter, I will employ my novel data set to address this question, by leveraging the information on unemployment status in the ACS, along with the subsequent employment outcomes in the LEHD data. This ability to observe the unemployment status of the worker before making their migration decision is advantageous compared to using the CPS or other sources that only observe migrators after the fact. Also, because the decision to change locations is endogenous to future employment outcomes, I will employ an instrumental variable technique that benefits from the precision afforded by the large size of my data set. Other longitudinal data sets, such as the NLSY, are typically too small for such estimation. In fact, the only prior study that attempts to systematically control for the endogeneity problem in a similar way uses a large administrative database from outside the U.S. (Pekkala and Tervo, 2002).

1.4 Creating Measures of Migration

In this section, I discuss how migration is measured in the existing data as well as in my new data source, and then compare some of the aggregate statistics derived from them. When measuring migration, the first task is to define what sort of geographic transition constitutes a “move” of interest. Because the theoretical framework describes migration as an attempt to expose oneself to different economic conditions that are superior to one’s present situation, economists are generally concerned with measuring moves that represent a substantial change of an individual’s local labor market.

Some standard definitions of labor markets have been created by government statistical agencies such as the concept of a Metropolitan Statistical Area (MSA), now known as a Core Based Statistical Area (CBSA), which groups collections of counties connected to a common urban center based on commuting patterns. This definition also proves useful for studies considering the housing market, as in my research here, since these boundaries also generally enclose the areas where the local workforce resides. Thus, a move to or from a CBSA likely represents a change in housing markets as well as labor markets. Several of the publically-available data sets described above are capable of providing CBSA-level information, although changing definitions of CBSA boundaries cause some complications for researchers. With county-level information, however, CBSAs can be reconstructed even if the data source does not explicitly report them. Other narrowly defined labor market areas include Economic Areas (EA), Public-Use Microdata Areas (PUMA), and Workforce Investment Areas (WIA), although these are generally less available in public data.

Because these definitions of labor markets do not cover the entire country, by generally excluding rural areas, any statistics aggregated to these levels will by definition be less than comprehensive. As such, many aggregated statistics report cross-state migration, thus capturing all migration within US borders in a way that is consistent over time. Despite the fact that some labor markets span across state borders, state-to-state moves also generally represent a substantial change in one's employment situation and economic conditions. Tracking moves across groups of states composing the four Census Regions also generates broad and meaningful measures of national migration. These types of definitions also provide the best point of comparison between different data sources, as nearly all data sets provide information at least at the state level, whereas they may not report narrower units of geography. Thus while I will use CBSAs as the basis for testing the hypotheses in my analytical chapters, later in this section I will also report some measures of cross-state and cross-region migration rates in order to compare my new data set with other sources. Because the LEHD microdata provides geographic detail about the establishment where an individual works down to the Census block, I can aggregate my measures of migration to virtually any geographic level.

Additionally, researchers must select the timespan over which to consider migration. This selection may obviously depend on the topic at hand, although different data sets offer more or less flexibility in terms of timing. In the Decennial Census, ACS, and CPS the time period is defined as either 1 year or 5 years prior to the survey date, with no further ability for the researcher to vary the timeframe. However, those surveys that include the individual's place of birth do allow for a measure of lifetime mobility or return migration, although place of birth does not necessarily denote a location where someone has spent an extended period of time.

Also note that with these retrospective questions it is possible that there have been multiple moves since the reference period, and thus some migration may go unobserved, especially in the case of the 5-year migration question. This concern is true for all measures based on two distinct start and end points, rather than a continuous longitudinal tracking.

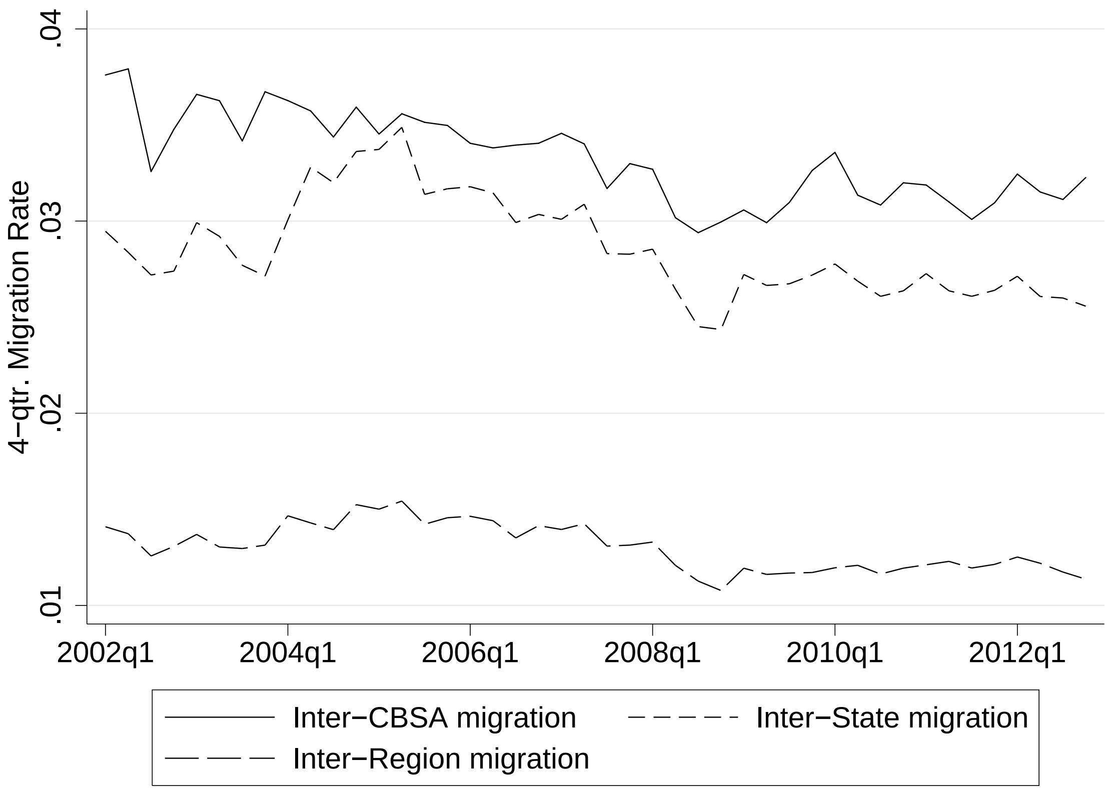
The longitudinal data sources, on the other hand, provide more options for selecting a time span over which to detect migration. The SIPP interviews participants every 4 months, allowing for a high-frequency definition of migration, although the same individuals are tracked for only 2-5 years depending on the particular panel. The NLSY79 tracks participants for a long period of time, but interviews are only conducted every two years as of 1994. The confidential IRS microdata allows the researcher to track the location of tax filers (not dependents) annually, but the lack of other covariates limits the possible analyses. My new data set, on the other hand, takes advantage of the quarterly frequency of the LEHD employment data, allowing for a higher-frequency measure of migration than other sources. Although designed as an annual survey, the ACS microdata also provides a specific date of interview, allowing the original time of observation to be placed in a specific month and quarter as well. Therefore, any individual appearing in the ACS at any point can be tracked quarterly in the LEHD data for all subsequent time periods.

1.4.1 Comparing Migration Rates Between Data Sources

Next we look at aggregate migration statistics calculated using some of the data sets described above. In particular, we will see how these rates have changed over time, and how the metrics from my new data source compare to those that are typically used. For reference, statistics from the other sources are taken from the work of Molloy et al. (2011), which provide a comprehensive overview of facts about migration in the United States.

The fraction of people who are observed relocating during the course of a year is small but significant, and has been found to be fairly consistent across different data sources. In order to make comparisons with existing measures, Figure 1.1 displays the annual migration rates as calculated from my new ACS-LEHD data set, for the time period representing 2002-2012. While the construction of the data will be described in more detail in subsequent chapters, annual migration is defined as being observed as newly employed in the LEHD system in a different location than where one was originally observed in the ACS, within four quarters of the interview date. The figure shows the average mobility rates for all individuals aged 18-65, weighted by the inverse propensity of appearing in the ACS.

Figure 1.1: Migration Rates from Combined ACS-LEHD Data



The rate at which individuals move across broad regions, (defined as the four Census Regions of Northeast, Midwest, South, and West), has ranged between 1.1 and 1.5% and has been declining over much of the sample period. While the rate appears to rise slightly over the first part of the decade, note that this coincides with a time in which new states were being added to the LEHD data, thus mechanically increasing the detection of migration[2]. By 2005, however, when nearly all states were included, the rate begins to nonetheless fall. The steepest period of decline occured between 2005-2008, coinciding with the housing crisis and beginning of the Great Recession, with some stabilization in the rates occuring thereafter. These magnitudes and features of the data are similar to those found by Molloy et al. (2011), especially to those derived from the ACS and IRS data. As noted by those authors, the CPS migration measures are consistently lower than in all other sources during recent years, although the precise reasons are unknown. While the time-series of the ACS-LEHD statistics only go back to 2002, the declining rate appears to be a continuation of the decline that has been observed in the CPS going back to at least 1980, with most measures also displaying a particularly steep decline during the Great Recession.

Inter-state migration rates have hovered around 3% on an annual basis, implying that about half of inter-state moves occur within the same region of the country. Inter-CBSA moves are somewhat higher at 3-4% annually, reflecting the fact that some individuals are observed changing labor markets within the same state. These

[2] Further description of the varying geographic coverage in the LEHD data, and the resulting measurement issues, is provided in Chapter 2.

numbers are roughly comparable to the corresponding measures in the ACS and IRS data, albeit higher by a couple tenths of a percentage point[3]. Note however, that while my sample measures the job mobility of individuals aged 18-65, other sources generally measure household mobility weighted by the number of residents, who may be younger or older than the 18-65 age range. In other words, it should perhaps not be surprising that a migration rate based on the observed job mobility of the working age population is somewhat higher than the residential mobility rate of the population at large.

In order to make more detailed comparisons between data sources, Table 1.1 reports the average annual migration rate for various demographic and socio-economic subgroups in the ACS-LEHD data set compared to the corresponding calculations from the CPS during the first decade of the 2000s. Due to the start date of the LEHD data, however, the decade is truncated by 1 year in my data set.

[3]For comparison, see Figure 2 in Molloy et al. (2010).

Table 1.1: Migration Rates by Subgroups: ACS-LEHD Dataset vs. CPS

	ACS-LEHD Dataset: 2002-2010	CPS: 2001-2010 (Molloy, Smith and Wozniak)
Sex:		
Female	2.5	1.6
Male	3.4	1.7
Age:		
18-24	2.5	3.0
25-44	1.8	2.2
45-64	1.4	1.0
Race:		
White	3.2	1.8
Black	2.8	1.7
Education:		
Less than high school	2.7	1.0
High School	2.4	1.2
Some college	2.8	1.5
College+	3.5	2.1
Employment Status:		
Employed	3.0	1.6
Unemployed	5.2	3.5
Not in Labor Force	2.3	1.5
Income:		
Bottom 50%	2.8	1.7
Top 50%	3.0	1.6
Children in the household:		
No children	3.1	2.0
At least one child	2.7	1.4
Homeownership:		
Renters	4.7	3.5
Owners	2.2	0.9

Notes: The left column reports mean 4-qtr inter-state migration rates for the given subgroup across the entire period in the combined ACS-LEHD data set, weighted by the ACS survey weight. The right column reports the corresponding 1-yr (retrospective) migration rate from the CPS, as calculated by Molloy, Smith and Wozniak (2011). All stratifying variables are based on similar definitions between the two surveys.

We see that while the rates calculated from the ACS-LEHD data set are uniformly higher than those in the CPS, due to reasons mentioned above, the patterns across the subgroups largely mirror one another in the two data sources. Demographic differences are evident in the table, with males being found to have a slightly higher mobility rate than females, and younger workers appearing far more mobile than those approaching retirement age. Also, whites are somewhat more mobile than African Americans, and the college educated are the most mobile educational group. Household factors are also significant, as households with children display less mobility. Additionally, there is a large discrepancy between homeowners and renters, with those who live in an owner-occupied home being far less likely to move. While this clearly conflates many demographic differences between owners and renters, measures like these have led researchers to investigate the hindering effects that homeownership may have on mobility, particularly in the context of the housing bust of the 2000s that left many owners underwater on their mortgages – thus motivating the analysis in the following chapter. There are also stark differences across employment statuses, with the unemployed displaying mobility rates nearly twice as high as for the employed. This is consistent with the notion that the unemployed use migration as a way to escape joblessness, also a topic that I address in a later chapter.

1.5 Conclusion

This chapter has discussed the different sources of data that are employed in the research of internal migration, and motivated the use of a novel data source that combines survey and administrative sources in order to address some shortcomings in the migration literature. In contrast to the existing sources of data, this combined ACS-LEHD data simultaneously provides a direct and flexible measurement of labor mobility, a large sample size, and the crucial ability to observe the conditions of individuals prior to the migration decision. This allows me to more accurately study the determinants of migration, such as homeownership, as well as study the consequences for certain segments of the population, such as the unemployed. The typical data sources used in the field are not as well-suited to addressing these topics, as they either detect migration retrospectively and thus do not record homeownerhsip and employment status at the origin location, or else are too small to provide a high degree of statistical precision.

Identifying the reasons for the recent decline in geographic mobility is very much an ongoing inquiry, as is the search into its ramifications for the labor market. The two analytical chapters that follow will use my newly created data set in order to test a pair of hypotheses that address these issues. In my first analysis, I find that a homeowner who has experienced a decline in the value in their home are substantially less likely to be observed starting a new job in a different location, compared to a comparable renter in the same place and time. I find some evidence of lock-in affecting the unemployed, although the implied impact on the unemployment rate is not likely to be significant. The final chapter asks whether the assumption that migration is beneficial for the unemployed is a good one, as little empirical data currently exists to support it. Indeed, my findings show that unemployed workers who move are able find employment in a shorter time frame than they would have had they remained, thus suggesting that mobility is indeed a valuable tool for such disadvantaged workers.

Whether or not migration is particularly important for underwater homeowners, the unemployed, or other vulnerable subgroups, the recent decline in geographic mobility should be viewed in the context of a widespread decline of labor market dynamics more generally, which may have serious and long-lasting effects on the economy. While the popular notion is that workers today switch their jobs more frequently than in the past, echoing the idea of the U.S. as a mobile nation, a growing body of data and research suggest that neither is the case (Molloy et al., 2011; Hyatt and Spletzer, 2013). This decline in fluidity appears to hold true on the firm side as well, with U.S. businesses exhibiting lower entry and exit rates over recent decades (Decker et al., 2014). While this may reflect a natural shift in the way resources are allocated in the modern economy, researchers have expressed concerns about the potential consequences for labor market search, wage increases, unemployment duration, as well as output and productivity growth (Davis and Haltiwanger, 2014; Hyatt and Spletzer, 2014; Molloy and Wozniak, 2011). No single factor has been found to explain much of the decline in these dynamics, although many have posited that they are interrelated (Molloy, Smith, and Wozniak, 2014). Thus, in exploring the issues that lead to and stem from migration, one must keep in mind that the movement of workers across the country is inextricably tied to the labor market as a whole. Exploring these complex interactions will obviously be a field of study for years to come, and it is clear that new and improved data sources, such as the one described here, will be useful in providing answers to these questions.

Chapter 2: Falling House Prices and Labor Mobility: Testing the Lock-In Hypothesis

2.1 Introduction

The mobility of labor has long been considered a hallmark strength of the U.S. economy, however, during the Great Recession the fraction of Americans that migrated between states fell to levels not seen in half a century (Molloy, Smith, and Wozniak, 2011). Many researchers and commentators have suggested that a driving force behind this decline was the historic collapse of the housing market, which left workers unable or unwilling to sell their homes. In this chapter, I explore this hypothesis that falling house prices deter American workers from relocating to different areas for new jobs, by effectively trapping them in their current residence. With nearly a third of all homeowners in 2009 owing more on their mortgages than their homes were worth, referred to as a being in "negative equity" or "underwater", and with millions remaining in this situation today, the consequences of the so-called "lock-in" effect for geographic mobility could indeed be dire.

Such an impact on migration could be harmful to the broader economy in multiple ways. First of all, if a crash in housing prices reduces out-migration particularly from the distressed areas that are hit hardest by the economic downturn, it could impede the efficient reallocation of labor and potentially exacerbate the type of "jobless recovery" that marked the years following the Great Recession[1]. In other words, the housing market may be serving as a friction that increases the baseline level of structural unemployment. However, besides the immediate job prospects for the unemployed, lock-in may also hinder the efficient reallocation of labor across the country in terms of overall job flows, by preventing the normal search process that leads to high quality job matches. As argued by Davis and Haltiwanger (2014) this decline in worker fluidity can in turn affect overall employment and productivy levels in the long run. Indeed, it is well documented that many types of these job-to-job flows decreased substantially during the Great Recession (Hyatt and McEntarfer, 2012), and deterred migration may be one factor contributing to this downturn.

Nevertheless, it is theoretically ambiguous whether a crash in housing prices should directly hinder labor mobility. In fact, one may even expect to see an increase in mobility as a consequence, because falling house prices and the resulting negative equity have been shown to be the driving forces behind home foreclosures (Gerardi, Shapiro and Willen, 2008). The estimated 11 million households that went into foreclosure between 2007-2010 represent a group of workers who were potentially more mobile due to the housing collapse[2]. Furthermore, lower prices do not, in and of themselves, provide a financial disincentive to relocating, since the diminished proceeds from selling one's home are mitigated by the lower prices that other homes can be bought for across the country. In fact, lower house prices can lessen some of transaction costs of moving, particularly those that are based on a percentage of a

[1]For an example of such concerns, see commentary by Lawrence Katz, *New York Times* Jan. 10, 2010.

[2]RealtyTrac, U.S. Foreclosure Market Report 2010.

home's value, such as broker fees and taxes.

There are two principal reasons, however, why researchers have generally hypothesized that labor mobility should fall as a result of declining home prices. First, homeowners who are underwater on their mortgages may avoid relocating if they do not have the liquidity to cover the monetary costs associated with moving and buying a new home, such as the required down payment, as well as taxes and transaction costs. The housing finance literature has long focused on the role of such financial constraints in the decision to sell one's home, developing models to explain how changes in house prices, interest rates, and down payment requirements affect turnover in the housing market (Quigley, 1987; Stein, 1995).

Secondly, even if such financial constraints are not binding, homeowners may be unwilling to move if it forces them to realize a capital loss on their homes (Genesove and Mayer, 1997). This notion of nominal loss aversion stems from the prospect theory literature pioneered by Kahnemann and Tversky (1979), who demonstrate how a utility function that is kinked or discontinuous at a "reference point" can cause individuals to react strongly to nominal losses. Thus, if an owner is unable sell their home for at least a certain amount, such as the original purchase price, they may be particularly unwilling to sell their home.

While several empirical papers have explored the dynamics between house prices and worker mobility, many data sources on migration only observe individuals after they have moved, thus preventing a direct measurement of the conditions contributing to geographic mobility. Furthermore, data that do allow researchers to track individuals over time are not recorded with sufficient scope and frequency to account for other local economic factors that may influence both migration and house prices. In this study, I will use the unique dataset introduced in Chapter 1, which combines a large survey of individuals with administrative data on their job histories across the country. This permits me to observe the homeownership status of the large sample of workers in the ACS, and follow their subsequent employment outcomes in the LEHD database in order to detect migration.

With these data I am able to measure the difference in migration propensities between homeowners who are in negative equity and those who are not, relative to a control group of renters who should be unaffected by changes in home values. By incorporating localized house price information from the real-estate data company Zillow, I test how a positive or negative change in home values since moving into one's residence differentially affects the out-migration propensities of homeowners relative to renters in the same metropolitan area and time period. This type of within-geography analysis will enable me to disentangle the influence of house prices from the impact of unobserved local economic factors that also affect migration behavior.

Specifically, using a within-CBSA and time regression specification, I find that a homeowner who has experienced a nominal house price decline is 0.2 percentage points less likely to migrate for new employment than a homeowner who has experienced price growth, relative to the difference between equivalent renters. This indicates that such a worker is approximately 12% less likely to relocate than the overall sample average of homeowners. Moroever, given that a third of homeowners experienced such a decline after the housing bust, the implied effect would account for more than 20% of the 0.3 percentage point decline in homeowner mobility between 2006-2010. Other specifications, which measure the impact of a given percentage change in home prices, show that a 25 percentage point drop in home prices would translate into a 0.13 percentage point drop in average homeowner mobility, accounting for 40% of the observed decline. Additionally I find evidence that these effects are also present in a subsample of unemployed workers, although they are mitigated somewhat and do not suggest a significant impact on the national unemployment rate.

The rest of this chapter is organized as follows. Section 2 discusses the previous evidence on housing lock-in, and motivates the use of my new data set for tackling this hypothesis in a novel way. Section 3 describes the data sources and the construction of the combined survey-administrative data set described above. Section 4 explains how I measure labor migration in this data set, and reports some of the aggregate migration statistics that are derived from it. Section 5 presents the methodological framework, which is based on regressions that exploit the unique within-CBSA and time variation. Section 6 presents the results of the analysis, including baseline estimates, results focusing on the the population of the unemployed, and a series of robustness analyses. Section 6 concludes.

2.2 Existing Literature and Motivation

The early empirical evidence on the lock-in phenomenon focused on the ramifications for the housing market itself, with Quigley (1987) finding that homeowners with favorable interest rate terms on their mortgages were less inclined to change residences, since doing so would increase their monthly loan payment. Genesove and Mayer (1997) was the first study to specifically consider the constraints resulting from a lack of home equity, and using a sample of home sellers in Boston found that those with high loan-to-value ratios tended to keep their houses on the market longer in order to hold out for higher selling prices. In a subsequent paper, Genesove and Mayer (2001) found that price changes also played an important role, with owners experiencing nominal losses being more reluctant to sell. Ferreira, Gyourko, and Tracy (2010) used measures of homeowner equity from the American Housing Survey (AHS), and also found that negative equity leads to lock-in, reducing the probability of moving from one's residence by about 50 percent.

Despite the substantial body of evidence on the presence of lock-in, however, it is important to note that this phenomenon does not necessarily translate into a commensurate impact on job mobility. After all, the above-mentioned studies employ housing unit-level data that do not follow an occupant after they move, and consequently do not directly address geographic mobility. The fraction of workers who are willing to migrate long distances for jobs at any given point in time is relatively small, and it is not known *a priori* whether this sub-population is more or less susceptible to the lock-in phenomenon than the average homeowner. It could also be the case that workers who are presented with a employment opportunity in another location are able to accept the new job without selling their house, for example, by renting their home or by commuting long-distance. For these reasons, in my empirical specification, I model the probability of a worker moving to a new job in a different location–not simply selling one's home–and estimate how house price movements affect this propensity. Only in this way can we directly measure whether the lock-in phenomenon has a significant effect on job mobility and the economy-wide flow of labor.

Few previous studies have used individual-level longitudinal data that allow the researcher to determine where an individual has moved to. Chan (2001) studies homeowners in New York City, and concludes that those who faced financial constraints due to low home equity were 25-33 percent less likely to move. On the other hand, Engelhardt (2003) finds some evidence that reduced home equity constrained the inter-metropolitan mobility of young homeowners in the National Longitudinal Survey of Youth between 1985-1996. Note that both of these studies are based on small samples, however, and occurred well before the most recent recession. Because my sample of ACS respondents from 2002-2012 is large and nationally representative, it offers an ideal setting to study lock-in during the historic collapse of housing prices.

More recent studies following the Great Recession have begun to cast doubt on the magnitude of the lock-in effect and its impact on labor mobility. Schulhofer-Wohl (2010) offered a rebuttal to Ferreira et al. (2010), arguing that the construction of the AHS data is biased against observing mobility of underwater homeowners[3]. Farber (2011) finds that the mobility of CPS and Displaced Worker Survey (DWS) respondents does not differ significantly for homeowners compared to renters.

[3]Ferreira, Gyourko and Tracy (2011) defend their original findings in light of the Schulhofer-Wohl (2010) critique.

Employing a similar strategy, Valetta (2012) finds that the differential in the unemployment durations between owners and renters in the CPS does not correlate geographically to the severity of the housing downturn. Other studies that employ such cross-geographic comparisons include Aaronson and Davis (2011) with the SIPP, Schmitt and Warner (2011) with the DWS, and Dennnet & Modestino (2012) with IRS data, with all of these studies failing to find a correlation between migration and regional house price changes.

Because migration is generally observed retrospectively in the data sources used in these recent studies, such as the CPS and IRS data, an individual's homeownership status in their original location is not known. This makes analyses based on comparisons of owners and renters difficult, although some studies have attempted to infer past homeownership status based on current status (Farber, 2011). Such assumptions could be problematic, however, if migrators tend to switch from owning to renting a home in a way that is correlated with their negative equity status. Additionally, as most of these prior studies are based on geographical differences in house price movements, it is difficult to systematically control for the corresponding differences in the local labor markets. This could bias cross-geographic analyses if house price movements are correlated with other local economic factors, as we may suspect they are.

Enhancing the ACS with the LEHD jobs data adds a longitudinal dimension that allows me to address several of the above-mentioned shortcomings in the previous literature. Because the ACS provides the individual's homeownership status in their original location, I am able to directly observe the conditions facing an individual when they chose to relocate, and accurately identify homeowners and renters. In addition, since the LEHD data includes the geographic location of each job, I can directly observe labor mobility in contrast to many studies which only measure residential mobility. Also, the size and scope of the sample provide me greater geographic, demographic, and temporal coverage than has been possible with most studies. Crucially, the large size of the ACS sample provides sufficient variation to conduct within-CBSA comparisons. The resulting difference-in-difference strategy compares the migration propensities between owners with negative and positive equity relative to a control group of renters located in the same place and time, thus controlling for local labor market conditions which may be correlated with housing prices.

2.3 Data Sources and Sample Construction

The sample used in this study comes from the American Community Survey, the largest annual survey of the American population conducted by the Census Bureau, interviewing approximately 3.5 million households annually. In my analysis I use all ACS respondents and spouses who are either homeowners or renters located in one of 356 metropolitan areas for which the data on house prices are available. Besides information on homeownership, the survey also provides data on how long the respondent has been living in their current residence. Addditionally, the survey contains a host of demographic and socio-economic variables. Since I am interested only in the effects of lock-in on labor mobility, I restrict the sample to those aged 25-

54. This prime-aged population excludes the youngest workers just finishing their education, as well as those approaching retirement, as these workers' locational decisions are likely to be motivated by different factors than those that we are considering here. The sample period spans the years 2002-2012, and the interview date of the respondents is aggregated to the quarterly level for analysis. The size of the ACS was smaller from 2002-2004 due to the survey being in beta status at that time, however the within-time nature of the analysis eliminates any bias from this feature of the data. After an ACS household is surveyed it is not surveyed again, so until now there has been little opportunity to determine anything about the workers in subsequent time periods.

It is for this purpose that I use the LEHD employment and earnings database, a collection of Unemployment Insurance (UI) records from 50 participating states plus the District of Columbia, which serves to generate statistics such as the Quarterly Workforce Indicators (QWI). The data are recorded at the worker-employer (i.e. "job") level, and report the earnings that each worker receives from a given employer during a particular quarter. These data are linked to information on the worker's employer establishment, sourced from the Quarterly Census of Employment and Wages (ES202) database at the Bureau of Labor Statistics. UI-covered employment represents over 95% of all private wage and salary civilian jobs and nearly all of the state and local government workforce[4]. Groups not covered by the database include the self-employed, the unemployed, and some government workers[5].

[4]For more information on the data and construction of the LEHD infrastructure, see Abowd et al. (2005).

[5]Workers with zero earnings in a given quarter are referred to as "non-employed", which encompasses the unemployed as well as those who are unobserved for other reasons.

With these data I can link an ACS individual via an individual identifier to all of their LEHD-covered jobs during the period 2002-2013. The crucial piece of information for the purpose of measuring migration is in the geographic location of the employee's workplace establishment. Although the workplace for a given individual is imputed in the case of multi-unit employers, an individual is assumed to remain at the same establishment for the length of their tenure with that firm. Additionally, the imputation is largely based on the proximity of establishments to an individual's residence. Both these features therefore serve to bias against the observation of spurious locational changes over time.

Geographic coverage of the LEHD data varies over the sample period, as more and more states began submitting data throughout the 2000s. While only 45 states are included at the beginning of my sample in 2002, the number rose to 49 by 2004, and 50 plus the Distric of Columbia starting in 2010. As a consequence, the fraction of U.S. private employment covered by the data grew from 87% in 2002 to over 95% by 2010 (Abowd and Vilhuber, 2010). Due to this, I limit the sample of ACS respondents to those who are residing in a state that is LEHD-covered at the time of the ACS interview. However, the incomplete coverage clearly affects the rate at which migration is detected, thus systematically suppressing the observed migration propensities, especially during the early part of the sample. Again, in the main regression analysis, the inclusion of time-specific geographical controls should account for the fact that individuals located in a particular place and time have a given set of destinations that they can be observed migrating to. However, robustness checks will also be used to examine the sensitivity of the results to the varying geographic coverage.

The data on home values come from estimates by Zillow, a real estate data analysis firm. Zillow's proprietary algorithm estimates the value of each home, and their published statistics provide the monthly median home value by zip code. This can be linked to the residential location of each ACS respondent via a crosswalk between the Census tract of residence and the Zip Code Tabulation Area (ZCTA). By combining this with ACS information on how long the individual workers have been living in their current residence, this enables a calculation of change in house prices that is relevant to each worker's tenure in their residence. This is important because a fall in house prices will affect an owner who bought their house recently more severely than one who bought long ago, since the longer tenured owners may have past years of home appreciation to buffer them from recent losses. Additionally, since the purchase price at the time the owner moved into the home may serve as the reference point in the context of nominal loss aversion, such a price change is expressed in relation to this point.

2.4 Defining Worker Migration

The hypothetical way by which lock-in can affect job migration is by preventing a current homeowner from taking a job located in a different area from their current location, because to do so would likely require incurring the costs of changing residences. Therefore, in order to study such out-migration, we must first define an appropriate unit of geography on which to base our analysis. For this purpose I employ the concept of a Consolidated Business Statistical Area (CBSA), formerly known as metropolitan statistical areas (MSA), as defined by the Office of Management and Budget and the Census Bureau. A CBSA is a cluster of contiguous counties surrounding at least one urban core, which are economically integrated with the core as measured by commuting ties[6]. Larger areas with cores of over 50,000 people are called metropolitan areas, while those with smaller cores of 10-50 thousand are known as micropolitan areas. This level of geography is well-suited for studying the interaction between housing markets and the out-migration of labor, since the CBSA is roughly defined as the area in which the workforce of a particular employment node also resides. Consequently, a worker who considers taking a new job outside of the CBSA must likely also consider the required change in residence, making this precisely the sort of job decision that is vulnerable to the influence of housing lock-in.

The outcome of interest in the following analysis is based on the concept of an out-migrator, a worker who leaves a particular metropolitan area in order to take a job in another location. The theory of lock-in predicts that a homeowner who has experienced house price depreciation will be less likely to transition to a new job in another area, because they are less willing or able to make the necessary change in residence than they would be otherwise. Thus, I define an out-migrator (or mover), $m_{ijt} = 1$, when the following criteria are satisfied:

1. An ACS respondent i is observed residing in CBSA j during quarter t

[6] See OMB publication 10-02, December 2009.

2. According to the LEHD database, individual i is no longer working in CBSA j in quarter $t+1$.

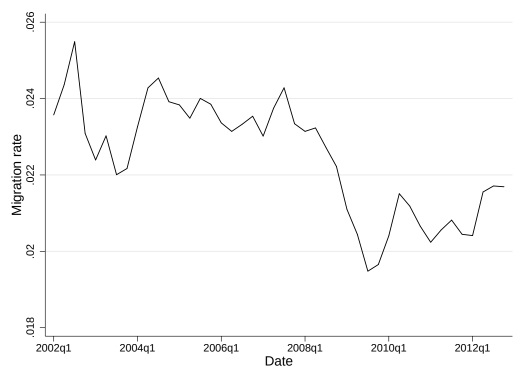
3. Individual i is observed beginning a new job (new employee-employer pair) at an establishment located in another area $-j$, within the threshold number of quarters of t.

On top of these three basic parameters, a few other qualifications must be made. For instance, if a new observed job located in $-j$ begins in the same quarter t as the ACS interview date, this job is required to last until at least $t+1$ in order to eliminate the possibility that this job in fact began before the ACS interview. Also, in case of multiple new jobs beginning at the same time, only the job with the highest earnings (or dominant job) is considered. Beginning a new job in a geographically adjacent CBSA that is considered to be connected to ones previous CBSA by commuting ties, defined as being part of the same Combined Statistical Area (CSA), is not counted as a move. Likewise, a new job located in a rural area of the same state will neither be considered to be a move. The 1-quarter threshold condition is meant to exclude cases where long periods of time have gone unaccounted for and therefore may not represent a clean employment transition, due to unobserved intervening unemployment or other non-covered employment. However, increased timing thresholds of two and four quarters will be also used in alternate specifications for robustness.

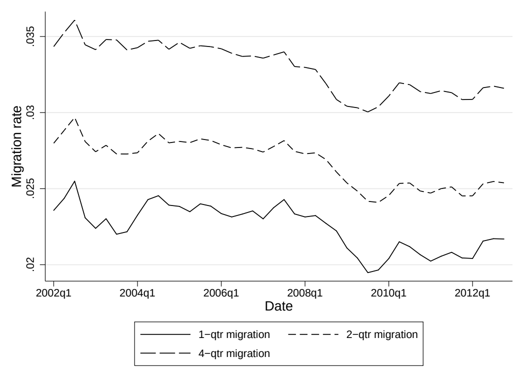
2.4.1 Aggregate Migration Rates

With this definition of out-migration in hand, we can measure aggregate CBSA-level migration rates by calculating $(\Sigma_{i=1}^{N} m_{ijt})/N$ over all workers i in quarter t. This measure expresses the propensity of the metropolitan workforce in time t to transition to a new job outside of the CBSA where a given worker is originally located. Figure 2.1 displays the time-series of the aggregate migration statistics during the sample period 2002q1-2012q3.

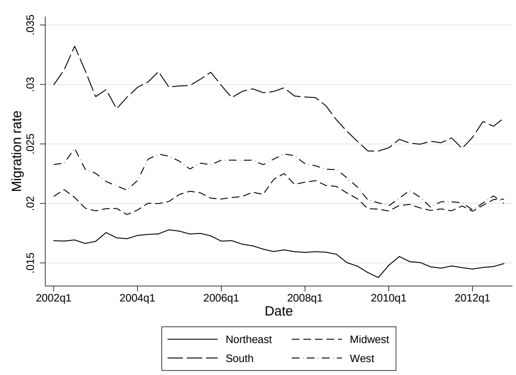
Figure 2.1: Overall Migration Rates



(a) 1-qtr Migration Rate



(b) Alternate Migration Thresholds



(c) 1-qtr Migration Rate by Census Region

Panel (a) shows the national migration rate for all individuals age 25-54, and we see that the migration rate declined throughout the decade, a feature of the data that has previously been noted in studies using other measures such as rates of interstate migration calculated using the CPS (Molloy et al., 2011). The quarterly CBSA-level out-migration rate stood at roughly 2.5% in the early part of the decade but had declined to below 2% by 2009:Q3. Also apparent is that the decline accelerated in the period between late-2007 and late-2009, corresponding with the housing crash as well as the Great Recession. Rates began to rebound by 2010, although only recovered about one third of their total losses by the end of the sample period in 2012. Out-migration rates using the 2 and 4 quarter thresholds reveal similar patterns, as seen in Panel (b). Finally, Panel (c) shows that regional variation exists as well, with the South Census region exhibiting the highest out-migration rates, and the Northeast the lowest. In terms of percentage change the West region saw the most dramatic decline, which coincides with the area of the country that was hit hardest by the housing bust.

Given the issues of LEHD coverage discussed above, the migration rates for the early part of the decade are less accurately measured because the set of destination states that individuals can be observed migrating to is reduced[7]. Furthermore, the missing states eventually joined LEHD at different times during the sample period, creating a somewhat inconsistent measure of migration. To assess the extent of this mismeasurement, Figure 2.2 plots the 1-qtr migration rate calculated exclusively

[7]States not included in the LEHD data as of the beginning of 2002 are: AZ, AR, DC, MA, ME, and NH.

from the set of 45 states that were consistently present throughout the entire sample, using only those same states as potential destinations for migration. As seen in the figure, this measure does not differ significantly from the measure using all states. While the discrepancy widens somewhat in the later years due to the changes in the underlying sample, the difference rarely exceeds a few hundredths of a percentage point, and the patterns mirror one another. Thus, due to these similarities, the entire sample of states will be used for the remainder of the analysis, although the consistent sample of states will be used later to test robustness.

Figure 2.2: 1-qtr Migration Rate With Consistent LEHD Coverage

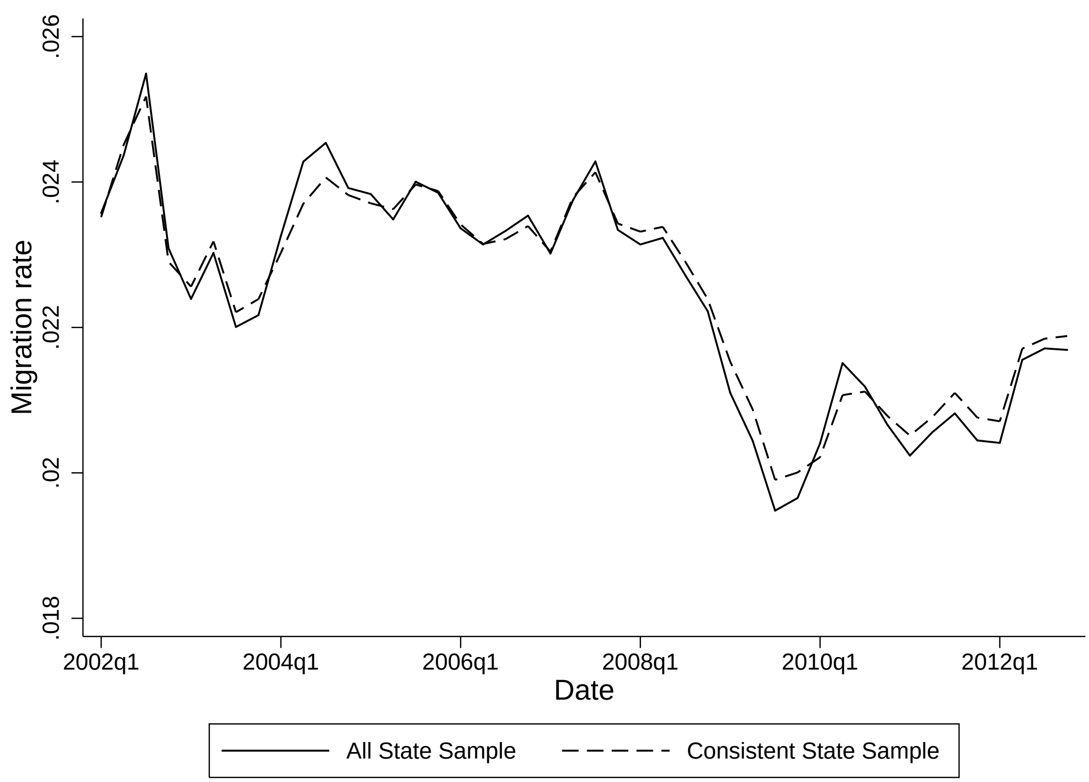


Table 2.1 reports the CBSA-level migration rates for various demographic and socio-economic groups. The patterns confirm some of the stylized facts about migration that have previously been demonstrated with other data sources, as discussed in Chapter 1.

Table 2.1: CBSA-Level Migration Rates by Demographic Characteristics

	1-qtr. Migration Rate	2-qtr Migration Rate	4-qtr Migration Rate
Sex:			
Female	2.0	2.5	3.1
Male	2.5	2.9	3.5
Age:			
25-35	3.0	3.7	4.5
36-45	2.2	2.5	3.1
46-55	1.7	2.1	2.6
Race:			
White	2.2	2.6	3.2
Non-white	2.4	2.8	3.5
Education:			
Less than high school	1.9	2.2	2.8
High School	2.3	2.8	3.4
Some college	2.4	2.9	3.5
College+	2.8	3.3	4.2
Employment Status:			
Employed	1.9	2.2	2.6
Unemployed	7.6	9.5	11.0
Not in Labor Force	2.3	3.1	4.4
Income:			
Top quartile	2.6	3.1	3.9
2nd quartile	1.9	2.2	2.6
3rd quartile	2.9	3.4	2.9
Bottom quartile	2.9	3.8	5.1
Marital Status:			
Married	2.1	2.5	3.1
Unmarried	2.6	3.1	3.9
Children in the household:			
No children	2.3	2.7	3.4
At least one child	2.2	2.6	3.2
Homeownership:			
Renters	3.3	3.9	4.9
Owners	1.7	2.1	2.6

Notes: Table reports percentage of demographic subsample that is observed out-migrating from their CBSA of residence according to a certain threshold. Demographic information is as reported on the ACS, and migration is subsequently observed in the LEHD employment data according to the above definition.

For example, migration rates are highest for the youngest age group, and decline monotonically with age. The higher educated also display high migration rates compared to those with less education. Males have relatively high migration rates, as do unmarried people and households without children. Figure 2.3 shows the time-series of migration rates for these demographic categories, and we see that there is some variation over time in the differences between these groups. In the later regression analysis, we must consider whether such changes in relative migration rates may be correlated with homeownership and price changes, because unobserved factors may have heteregeneous effects on these demographic groups and thus confound the analysis.

Figure 2.3: Migration Rates by Demographic Groups

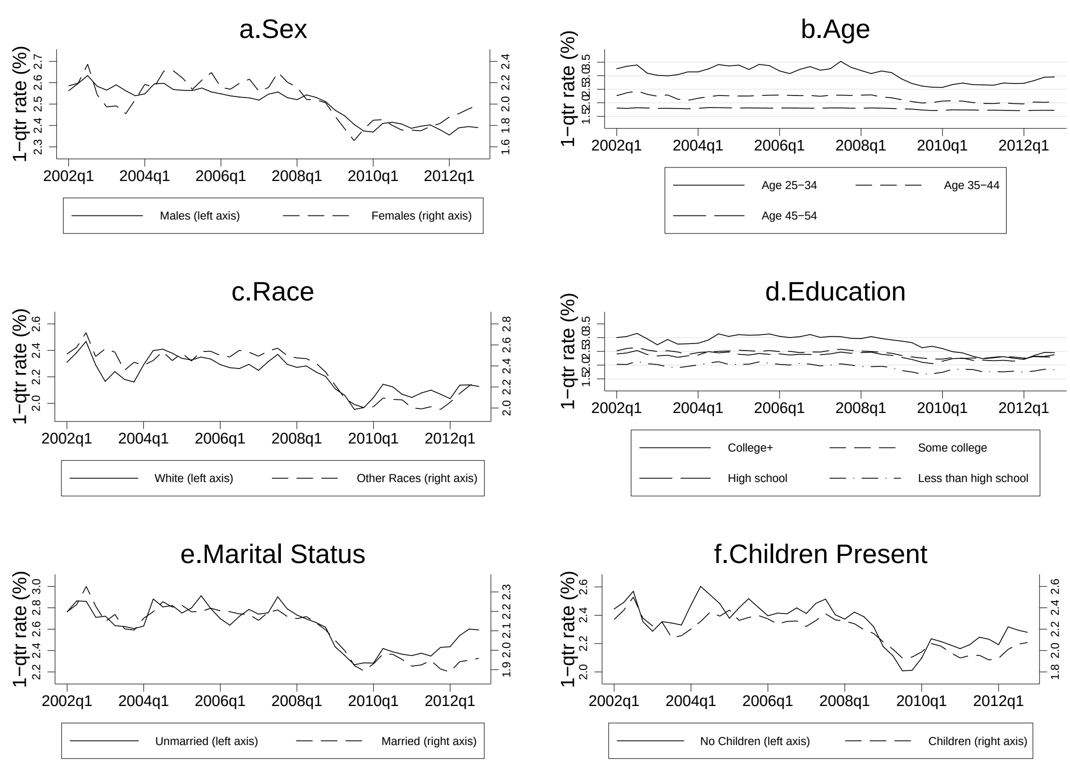
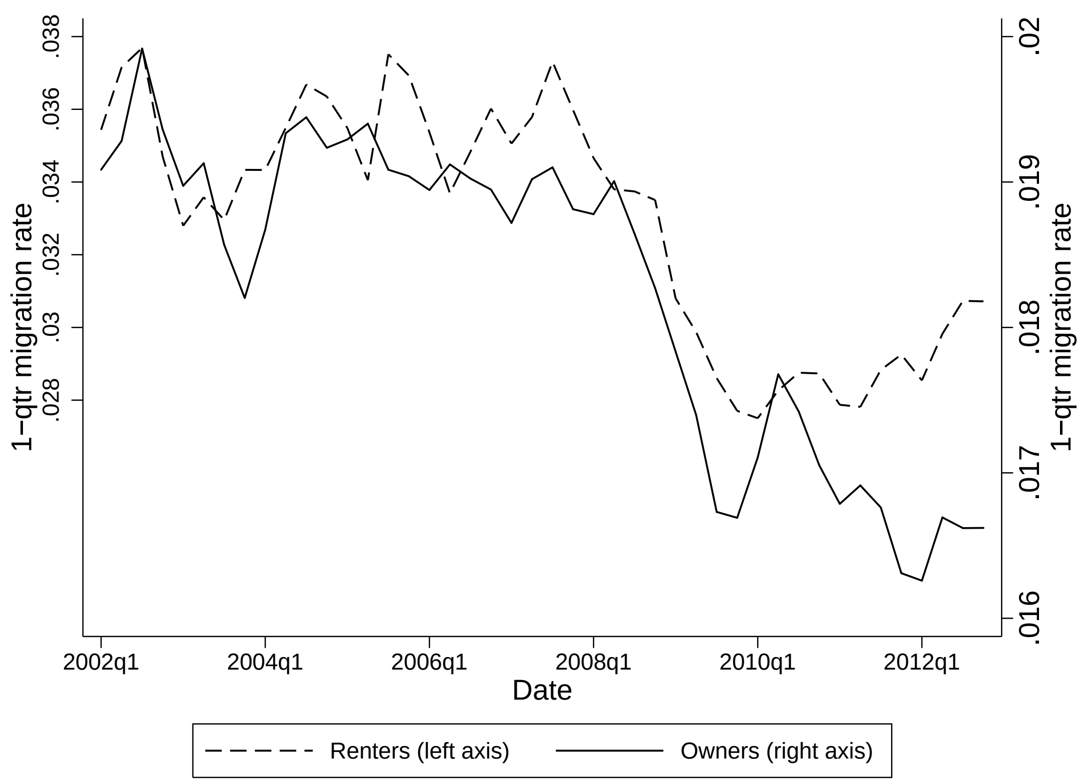


Figure 2.4 depicts the time-series of metropolitan out-migration separately for owners and renters on the national level. As seen back in 2.10, the mean migration rate for renters is 3.3% across the sample period compared to 1.7% for owners, thus the two lines are plotted here on different scales. Note, however, that this large mean difference is due partially to demographic differences, which are not controlled for here in the unconditional rates. However, comparing how the propensities of these two groups change over time can be instructive, and mirrors the strategy of the regression analysis that follows.

Figure 2.4: Migration Rates by Homeownership Status



The figure reveals that the two time series tracked eachother quite closely in the early part of the decade, but began to diverge by late-2005 when the migration rate of owners began to trend downwards somewhat. The migration rate for renters remains relatively steady until 2007 when the rates for both groups began to plunge. However, in 2010 it appears that the mobility rates for renters stabilized and and even started to rise, whereas the rate for owners has continued to drift slightly downward. Such differences in the mobility of owners and renters have not been noted in other studies using different data sources, although few such studies have used data through 2012. While this evidence is merely suggestive, we now turn to within-CBSA-time regressions in order to systematically control for differences in individual characteristics and confounding labor market factors.

2.5 Within-CBSA and Time Estimation

The goal of the key regressions that follow is to isolate the impact of house price changes on the out-migration propensity of homeowners, using the concept of migration defined above. To properly account for the many possible confounding factors that vary at the CBSA-time level, the estimation strategy uses individual-level data that vary by homeownership status, zip-code level house price exposure, and time since moving into their home. This methodology allows me to conduct a within-CBSA and time analysis that controls for CBSA-specific time shocks by comparing the response of homeowners to house price changes relative to a comparable set of renters located in the same CBSA and time period. Since experiencing a decline in house prices should not directly impact the migration decision of those who are renting, the differential impact of house price changes on owners relative to renters will be representative of the direct effect on owners. This use of renters as a comparison group, a technique used frequently in the literature, amounts essentially to the following differences-in-differences comparison:

$$[P(migrate/Owner_{jt\Delta HousePrice<0}) - P(migrate/Owner_{jt\Delta HousePrice>0})] \quad (2.1)$$
$$-[P(migrate/Renter_{jt\Delta HousePrice<0}) - P(migrate/Renter_{jt\Delta HousePrice>0})]$$

where both Owners and Renters living in the same CBSA j and time t are distinguished by whether they have experienced either negative price appreciation, $\Delta HousePrice < 0$, or positive price appreciation, $\Delta HousePrice > 0$, since moving into their homes. If this term is negative, it means that the impact of experiencing a price decline is greater for owners than for renters, thus reflecting the direct effects of the loss of equity. The unobserved labor market factors in CBSA j at time t, which may be correlated with house prices, affect both groups equally and will therefore be partialed out with the inclusion of CBSA*time level dummies. Similarly, any general correlation between house prices and migration that is common to owners and renters will also be differenced out by these CBSA-time controls.

Note that renters can indeed be affected by factors that are reflected in house prices, such as amenity values, including neighborhood quality, crime, public services, environmental factors, and other factors pertaining to the microeconomy of a local area. The impact of amenities on migration has long been studied in the migration literature since Thiebout (1952), as has their effect on property values (Kain and Quigley, 1970; Cheshire and Sheppard, 1995). Thus while not strictly a control group, renters are assumed to be equally affected by these indirect factors as owners, allowing us to ascribe any differential impact to the additional influence of changes in home equity on owners.

The individual-level data allow me to include individualized measures of housing price change, as well as personal characteristics that may also be correlated with migration and housing related factors. Therefore, the resulting regression specification takes the general form:

$$P(m_{ijt} = 1) = \beta_1' HousePrices_{it} * Own_{ijt} + \beta_2' Own_{ijt} + \beta_3' HousePrices_{it} + \gamma_{jt} + OtherControls + \varepsilon_{ijt} \quad (2.2)$$

where $HousePrices_{it}$ is a variable expressing the individual's change in house prices. The slate of dummy variables, γ_{jt}, ensures that identification is based solely on within-CBSA and time variation.

The dependent variable is the individual probability of out-migrating, $P(m_{ijt})$, where m is a $\{0, 1\}$ dummy indicating whether the ACS individual i, observed during time t, switches to a job in a location other than his current CBSA of residence j, according to the criteria stated in the section above. The threshold used to define migration will typically be the 1-quarter definition in most specifications, but the 2- and 4-quarter thresholds will also be used for robustness. The key explanatory variable in the analysis is an interaction term between the individual's change in house prices and an indicator of whether an individual is living in an owner-occupied home or not. Thus we let $Own_{ijt} = 1$ if individual i lives in an owner-occupied house in CBSA j in time t and 0 if they live in a rented home, according to the ACS. This interaction term thus expresses the differential impact of house price on owners, and reflects the direct effect of changes in home equity.

2.5.1 House Price Variables

The construction of the house price variables is meant to express the change in prices that the individual has experienced since moving into his or her home, which proxies for a gain or loss in home equity. Letting t_i^0 be the move-in quarter of person i according to the ACS response, I can estimate a percentage change in house price relevant to an individual's experience as:

$$\Delta ln(HPI_{it}) = ln(HPI_{z_j t}) - ln(HPI_{z_j t_i^0})$$

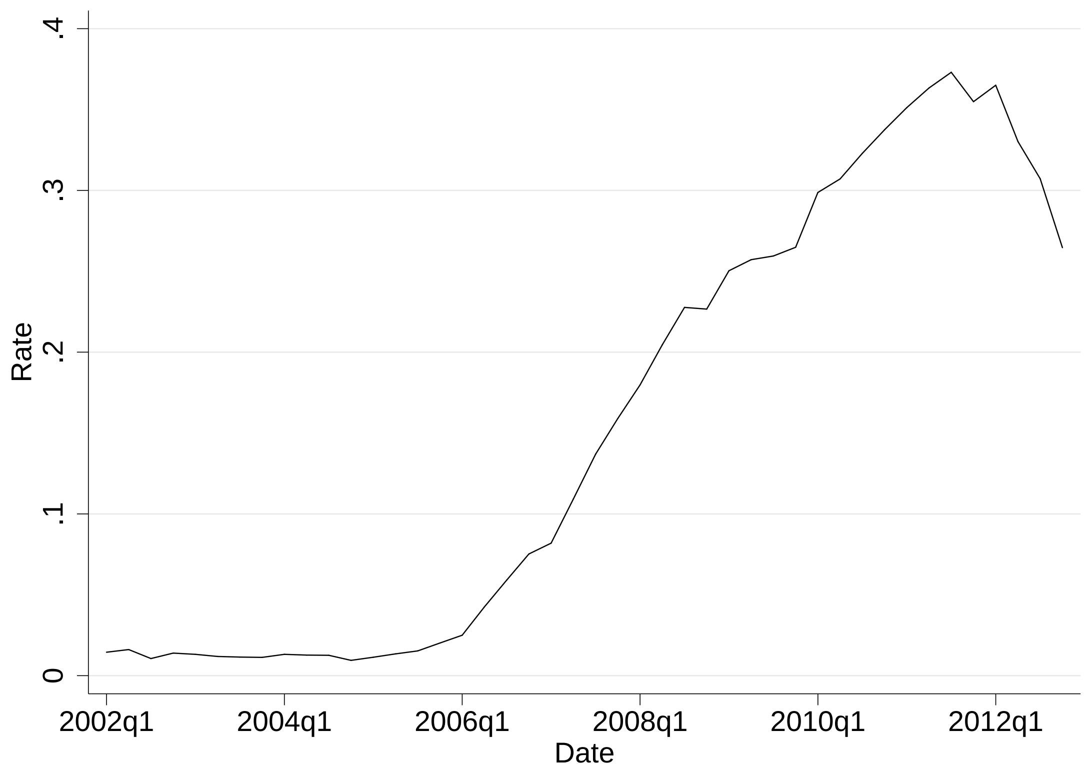
where $HPI_{z_j t}$ is the Zillow house price index value for person i located in zip-code z (within CBSA j) at time t, and $HPI_{z_j t_i^0}$ is the index at time t_i^0. In turn we also define:

$$neg\_equity_{it} \text{ if } \Delta ln(HPI_{it}) < 0, \text{otherwise } 0$$

which expresses whether an individual has experienced a home price decline since moving in to their residence, and can be interpreted as a proxy for negative equity amongst homeowners. As Zillow data only goes back to the late 1990's for most metropolitan areas, any individuals who moved into their home prior to the start date of the data are assigned the earliest available figure for that zip-code. There is some variation in when Zillow begins to report price data for different CBSAs, but the within-CBSA nature of the analysis will account for these discrepancies. Also, because the ACS does not define geography based on postal zip codes, I determine the zip code of residence using a crosswalk between an individual's census tract of residence and the corresponding Zip Code Tabulation Area (ZCTA), as defined by the Census Bureau.

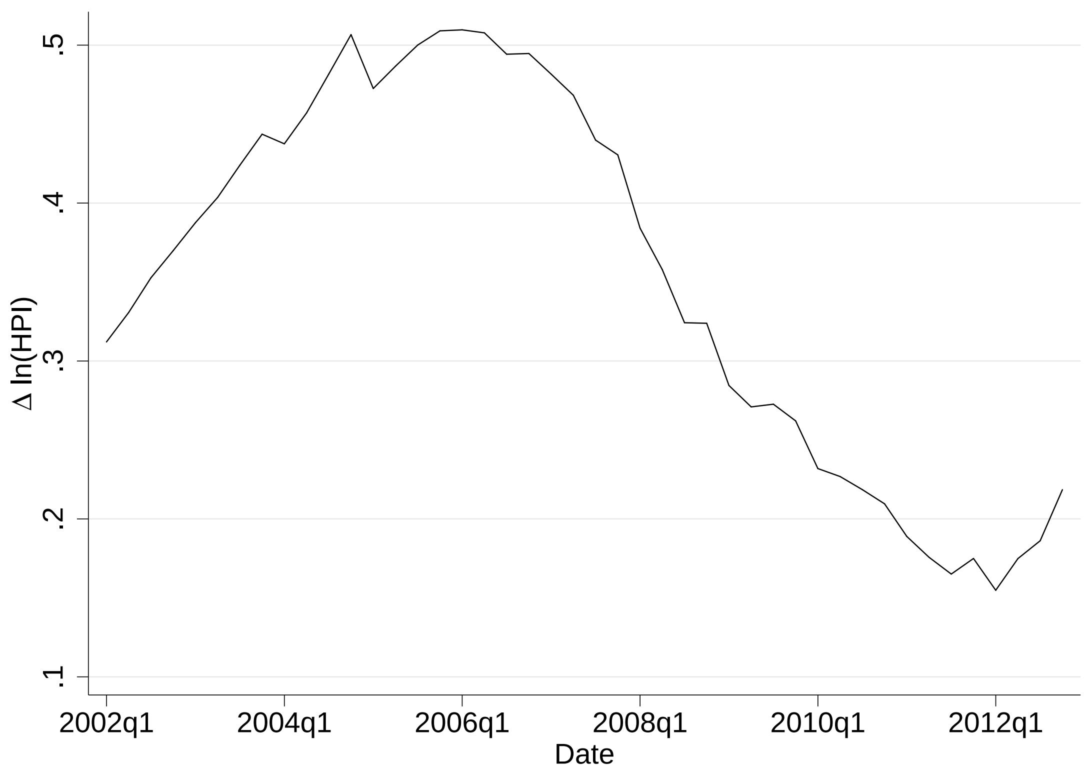
While house price declines are a necessary condition for negative equity, this estimate of negative equity status is of course inexact since we have no information on other things affecting home equity such as downpayments, second-mortgages and home equity loans. Estimates from the housing bubble years of 2001-2005 suggest that the median loan-to-value ratio at the time of mortgage origination was close to 90% (Demyanyk and Van Hemert, 2011). Thus, the proxy will likely overestimate the incidence of negative equity although for reasons discussed above, recall that a nominal price decline may in itself be a deterrent to selling one's home due to loss aversion. Figure 2.5 shows the "negative equity" rate for the entire sample of owners and renters, signifying the percent of the population that has experienced a decline in the value of their home during their tenure.

Figure 2.5: Negative Equity Proxy Rate



Due to a strong housing market in the 90s and early 2000s, this rate began virtually at zero in the early years of the sample, but began to rise in late 2006 until late 2010 when it reached around 35% nationally. During the last two years of the sample, however, the rate began to decline dramatically as house prices began to recover. Similarly, Figure 2.6 shows the mean home price appreciation over time, which rose from approximately 30% at the beginning of the sample to over 50% in 2006, before tumbling down below 20% by 2011.

Figure 2.6: Mean House Price Changes



In order to visualize the variation in the home price data, Figure 2.7 shows the distribution of the individual home price changes, as well as the residual values after regressing the price change variable on the γ_{jt}, $Own * \gamma_j$ and $Own * \gamma_t$ dummies. Due to the differences in prices on the zip-code level, and in the lengths of time that individuals spend in their homes, a significant amount of variation remains. There is also substantial variation over time, both before and after the housing crash. Figure 2.8 shows that the distribution was highly skewed rightward before the peak of the market in 2006q3, but became much more symmetrical between gains and losses after the decline.

Figure 2.7: Variation in Home Price Changes: Overall and Residual

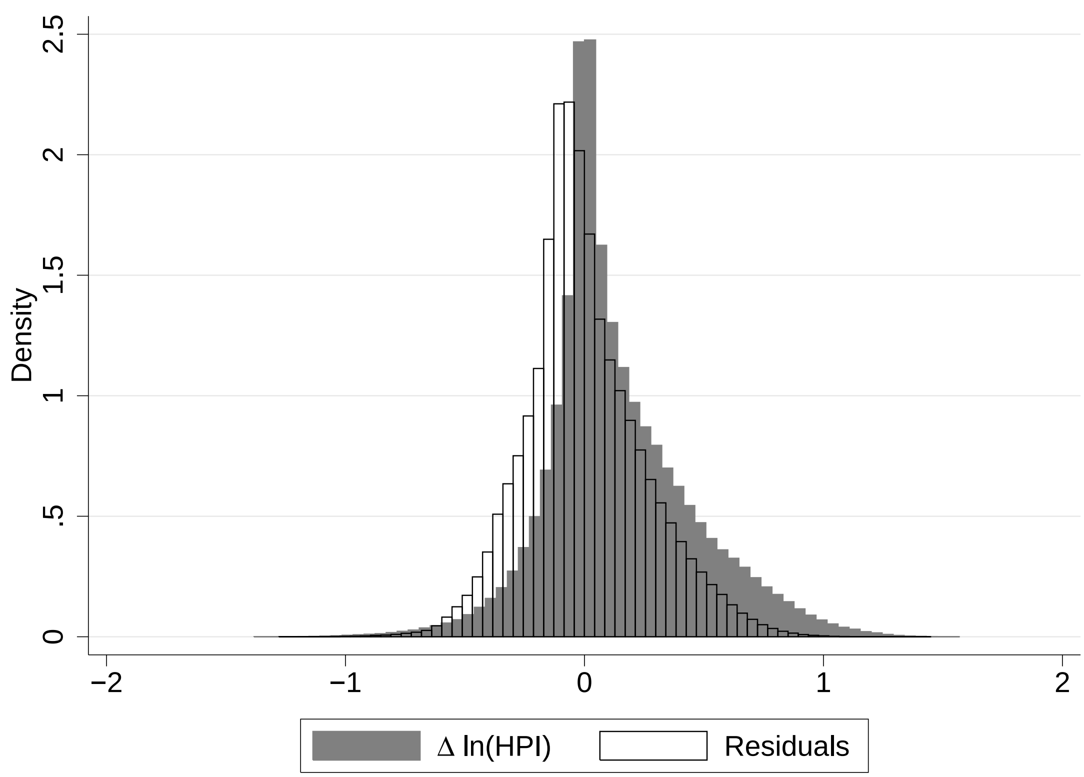
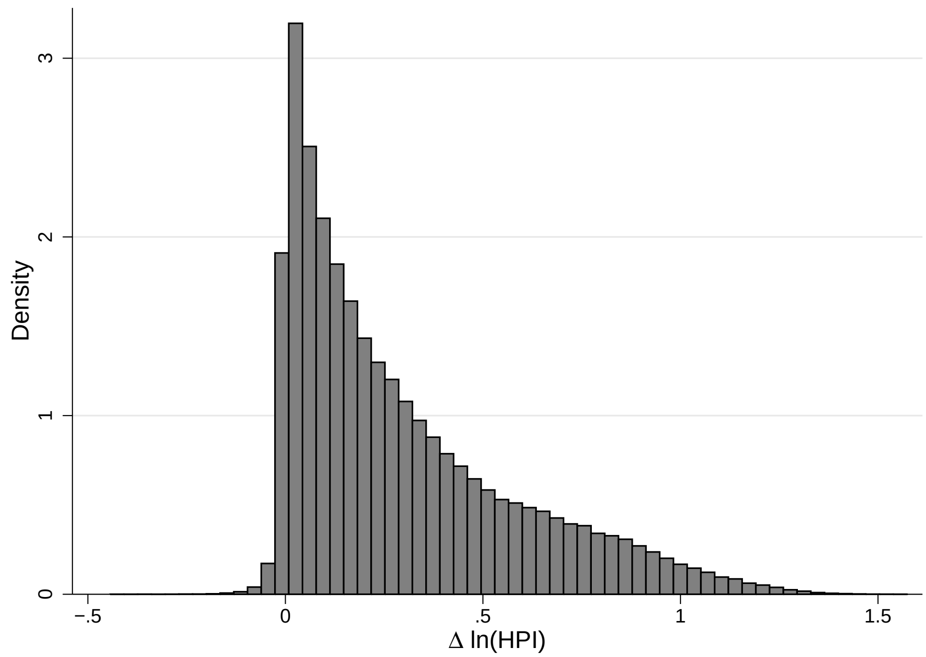
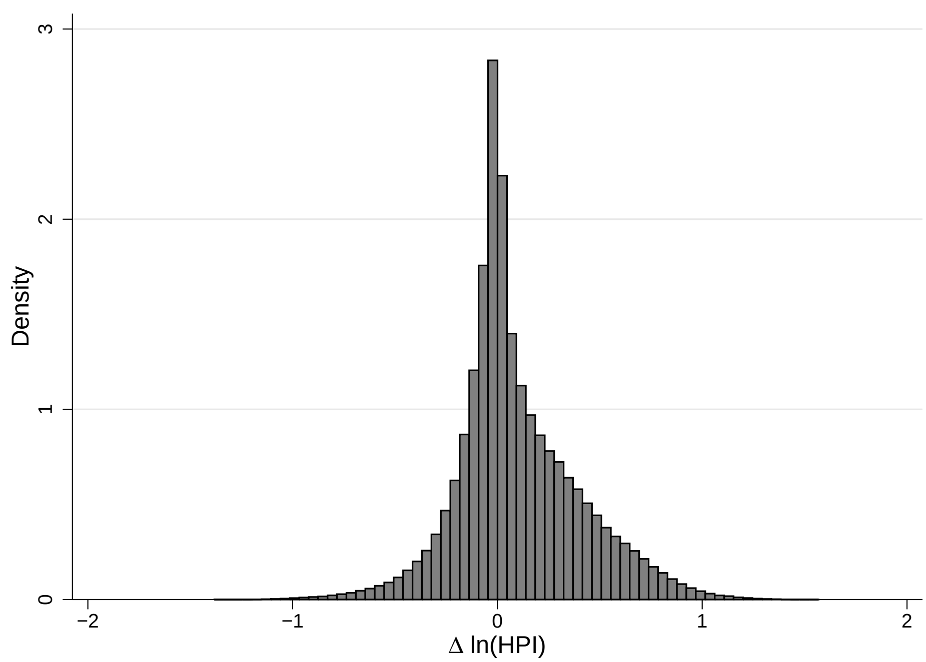


Figure 2.8: Variation in Home Price Changes: Pre- and Post- 2006q3



(a) Pre-2006q3



(b) Post-2006q3

Thus, given these two measures of house price changes, the crucial independent variable in the regression is denoted as $HousePrices_{it} * Own_{ijt}$, where $HousePrices_{it}$ equals either $\Delta ln(HPI_{it})$ or neg_equity_{it} depending on the specification. The coefficient on this variable can be interpreted as the differential effect of house prices on the out-migration decision on homeowners, relative to renters.

2.5.2 Control Variables

A variety of controls are used to ensure identification. First the crucial slate of CBSA-quarter level dummies, γ_{jt}, will control for CBSA-specific, time-variant shocks such as the local labor market conditions that affect buyers and renters equally. Two more sets of dummies, $Own * \gamma_j$ and $Own * \gamma_t$, will control for the average migration propensities of homeowners by CBSA or quarterly time period, respectively. These account for the fact that the migration propensity of owners is naturally different from that of renters, and this difference may vary by labor market, or over time on a national level. Note that since there are three levels of variation–CBSA, time, and ownership status–all three pairwise interaction terms can be included. Another key control accounts for the length of time that an individual has been living in their current residence, represented by a slate of yearly dummies ϕ_d (where d=1 through 49 years, and 50+). This is important both because housing tenure duration is likely a strong indicator of one's willingness to move, and also because the individual-specific house price variables are related to tenure length by construction. Finally, I include the individual-level X_{it} variables for age, sex, race, education, earnings quartile, marital status and presence of children, all as reported in the ACS.

Identification rests on the assumption that the homeownership decision is exogenous to unobserved factors that might be correlated with both house prices and migration propensities. Buyers and renters are clearly different even beyond their demographic compositions, due to factors related to self-selection, therefore their migration propensities will naturally differ. However, the previously described dummy variables control for the different mean migration propensities between the two groups, and these mean differences are even allowed to vary by CBSA ($Own*\gamma_j$) as well as by time nationally ($Own * \gamma_t$). Identification will only be threatened if there are unobserved factors that affect the *changes* in the relative migration propensities of owners and renters in a way that is correlated with changes in house prices. For instance, if house prices changes are correlated with something that has heterogeneous effects on migration for people with different characteristics, and these characteristics are strong determinants of homeownership status, it could lead us to incorrectly conclude that the migration of owners is being directly impacted by house price changes. Thus later we will explore possible heterogeneity in a series of robustness checks. Absent such confounding factors, however, if the respective migration rates of homeowners and renters react differently to house price movements, (after accounting for observable differences), it is due to the direct effect of prices on owners, in particular through the home equity channel.

Given the above definitions of all the variables, we can now write the complete regression specification:

$$P(m_{ijt}=1) = \alpha + \beta_1' neg\_equity_{it} * Own_{ijt} + \beta_2' Own_{ijt} + \beta_3' neg\_equity_{it} \qquad (2.3)$$
$$+ Own_{ijt} * \gamma_j + Own_{ijt} * \gamma_t + \phi_d + \theta' X_{it}$$
$$+ \gamma_{jt} + \varepsilon_{ijt}$$

Here, β_1 expresses the magnitude of this differential, since it represents the interactive effect between homeownership and the proxy for whether a worker has experienced declines in home prices during their tenure. A negative β_1 would indicate that the outmigration of owners is more negatively correlated with the negative equity proxy than the outmigration of renters, supporting the hypothesis that a decline in home equity decreases the propensity of homeowners to out-migrate.

In alternative specifications, we will use an independent variable expressing the log change in home prices instead of the negative equity dummy. The resulting equation is thus:

$$P(m_{ijt}=1) = \alpha + \beta_1' \Delta ln(HPI_{it}) * Own_{ijt} + \beta_2' Own_{ijt} + \beta_3' \Delta ln(HPI_{it}) \qquad (2.4)$$
$$+ Own_{ijt} * \gamma_j + Own_{ijt} * \gamma_t + \phi_d + \theta' X_{it}$$
$$+ \gamma_{jt} + \varepsilon_{ijt}$$

In this case, β_1 expresses the magnitude of the differential impact of a given percentage change in home prices on the migration propensities of owners. Here a

$\beta_1 > 0$ means that the outmigration of owners is more positively correlated with house price changes relative to renters, indicating the direct impact of changes in home equity on owners. In other words, it is a positive sign on the key interaction term that would support the lock-in hypothesis.

A linear probability model (LPM) is used to to specify $P(m_{ijt} = 1)$. While Probit and logit are other natural candidates for such estimation, the LPM is more straightforward for interpreting the marginal effects[8]. While index models are advantageous in that they restrict the support of the probability space to the unit interval, the inclusion of the large number of binary variables renders the estimation computationally onerous, as well as asymptotically inconsistent (Cameron and Trivedi, 2005).

2.6 Results

2.6.1 Baseline Results

The first estimation uses the whole ACS prime-aged worker sample to test for the lock-in phenomenon. As described above, this baseline specification is designed to estimate whether the measures of the change in house prices are associated with a decrease in a homeowner's probability of beginning a new job in another location, relative to the effects on similar renters. For these baseline regressions, recall that the sample of workers is composed of all ACS reference-persons and their spouses living in the same home, both homeowners and renters, restricted to be of prime working

[8]When a Probit is run on the baseline specification in equation 2.3, results are found to be qualitatively similar to the LPM.

age and located in one of 356 CBSAs. Standard errors in this specification, and all following tests, will be clustered by CBSA to control for the broad range of error dependencies that can occur between individuals residing in the same CBSA. Each regression is also weighted by the ACS survey weight, called the "person weight", which expresses the inverse of the probability of a given individual being sampled. Summary statistics for the baseline regression sample are shown in Table 2.2.

Table 2.2: Summary Statistics

Variable	**Mean**	**Std. Dev.**
$Migrate\_1q$	0.021	0.142
$Migrate\_2q$	0.024	0.155
$Migrate\_4q$	0.03	0.171
Own	0.76	0.427
$\Delta ln(HPI)$	0.31	0.367
Neg_Equity	0.216	0.412
$Female$	0.536	0.499
$Non-white$	0.192	0.394
$Married$	0.722	0.448
$Children$	0.439	0.496
$Age25-34$	0.226	0.418
$Age35-44$	0.264	0.441
$Age45-54$	0.223	0.416
$Less\_than\_High\_School$	0.081	0.274
$High\_School$	0.235	0.424
$Some\_College$	0.313	0.464
$College+$	0.37	0.483
$Earn\_quartile\_1$	0.25	0.421
$Earn\_quartile\_2$	0.25	0.421
$Earn\_quartile\_3$	0.25	0.421
$Earn\_quartile\_4$	0.25	0.421
$Employed$	0.762	0.426
$Unemployed$	0.042	0.2
$Not\_in\_Labor\_Force$	0.196	0.397
$\mathbf{N} = 7,300,000$		

Results from these baseline specifications are shown in Table 2.3. Column 1 reports the results from using the negative equity proxy, and the coefficient of -.002 on the interaction with homeownership reflects the differential impact of negative equity, compared to renters. This implies that experiencing a nominal house price decline is associated with a 0.2 percentage point decrease in the outmigration rate of owners with negative equity compared to those with positive equity. With a sample average migration rate of 1.7% for homeowners, this implies that a underwater homeowner is approximately 12% less likely to relocate than average. Given the precipitate decline in migration during 2006-2010 when the quarterly migration rate of homeowners fell by around 0.3 percentage points, and the fact that around one third of homeowners were in negative equity during 2010, one can estimate that the 0.2 percentage point differential accounts for about 0.067 percentage points, or over 20% of the total decline in homeowner mobility during this time.

Table 2.3: Baseline Results: All Workers

House Price measure	(1) Neg_Equity	(2) Δln(HPI)
	β/(s.e.)	β/(s.e.)
HousePriceXOwn	**-0.002*****	**0.005*****
	(0.001)	**(0.001)**
Own	-0.004*	-0.006**
	(0.002)	(0.002)
HousePrice	0.000	-0.002**
	(0.001)	(0.001)
Female	-0.007***	-0.007***
	(0.001)	(0.000)
$Non-white$	-0.000	-0.000
	(0.000)	(0.000)
Married	-0.001***	-0.001***
	(0.000)	(0.000)
Children	-0.002***	-0.002***
	(0.001)	(0.000)
Unemployed	0.055***	0.055***
	(0.004)	(0.004)
Not_in_labor_force	0.006***	0.006***
	(0.001)	(0.001)
$Age35-44$	-0.003***	-0.003***
	(0.001)	(0.000)
$Age45-54$	-0.006***	-0.006***
	(0.000)	(0.001)
High_school	-0.001**	-0.002**
	(0.001)	(0.001)
Some_college	-0.001	-0.001
	(0.001)	(0.001)
College+	-0.000	-0.000
	(0.001)	(0.001)
Earn_quartile_2	-0.001	-0.001
	(0.001)	(0.001)
Earn_quartile_3	-0.007***	-0.007***
	(0.001)	(0.002)
Earn_quartile_4	-0.006***	-0.006***
	(0.001)	(0.001)
N	~7,300,000	~7,300,000

* p ≤ 0.05, ** p ≤ 0.01, *** p ≤ 0.001

Notes: Sample consists of approximately 7.3 million ACS individuals i, either household head or spouse age 25-54, who are observed in one of 356 CBSAs j during a given quarter t in 2002q1-2012q4. All regressions include the following sets of controls: Ownership status X CBSA, CBSA X time period, Ownership status X time period (results omitted for clarity). Individual controls for sex, age, race, education, marital status, presence of children, and residential tenure length are also included. Standard errors are clustered by CBSA.

Column 2 reports results from the specification using the log change in house prices as the independent variable, and the coefficient on the key interaction term is 0.005, also consistent with the lock-in hypothesis. For interpretation, recall from Figure 2.6 that the mean total price appreciation of homeowners nationally fell to about 25% in 2010, down from 50% in 2006. Thus, this 25 percentage point drop can account for a decrease in the migration rate of about 0.125 percentage points, which represents about 40% of the decline in the migration rate of homeowners during this time. Note that this estimate represents an impact across the entire spectrum of both positive and negative price changes, and not only on those with negative equity, implying that the greater the increase in home equity the higher the willingness to undertake a long-distance job move. This may reflect that the threshold amount of equity required for moving differs by person, or that migration generally becomes more attractive as wealth increases[9].

Looking at the other coefficients in Table 2.3, we see that homeownership itself lowers the probability of migrating, with a magnitude that is bit higher than on the key interaction term but less statistically significant. In other words, homeowners are simply less mobile naturally due to self-selection, as well as due to the influences of changes in home equity. The home price variables themselves, however, do not appear to have much common impact on owners and renters, with a coefficient of essentially zero in Column 1 and a slightly negative coefficient in Column 2. This means that rising prices due to changes in things such as amenity values make both

[9]A regression that includes both the negative equity proxy and the log change in house price specification yields qualititatively similar results to the baseline, although the negative equity proxy interaction loses its significance.

owners and renters somewhat less likely to move away. Finally, as for the individual-level characteristics, we see higher migration propensities for the unemployed and lower earning workers, younger workers, males, the unmarried and those without children, largely consistent with previous findings.

The large magnitudes of the estimates stand in contrast to much of the recent literature, which typically have found little to no evidence of the lock-in effect. However, my estimates fall in line with much of the earlier literature which found that negative equity reduced household mobility by roughly 10-30% (Engelhart, 2003; Ferreira et al., 2011). Note also that no previous studies have employed the same within-CBSA and time strategy used here, instead relying on broader sources of variation. To see how different levels of variation affect the results, Table 2.4 shows estimates from the baseline negative equity specification by sequentially adding different sets of control variables.

Table 2.4: Baseline Results: Impact of Controls

Specification	**(1)** **No Controls**	**(2)** **Add Individual Covariates**	**(3)** **Add CBSA and time FE**	**(4)** **Add CBSA*time interactions**
	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)
Neg_EquityXOwn	**-0.001**	**-0.000**	**0.001**	**-0.002*****
	(0.003)	**(0.000)**	**(0.001)**	**(0.001)**

Notes: Sample consists of approximately 7.3 million ACS individuals i, either household head or spouse age 25-54, who are observed in one of 356 CBSAs j during a given quarter t in 2002q1-2012q4. The set of controls used differ by specification, as described in each column. Standard errors are clustered by CBSA.

Column 1 reports the results from a specification with no controls, and reveals no relationship between the negative equity interaction and migration propensity. Furthermore, adding individual controls in Column 2 does not notably change the estimates. Thus, these specifications exploiting wide geographic variation in house prices yield results similar to previous studies that are based on cross-CBSA or -state comparisons. Column 3 shows the results from adding CBSA and time controls, and interestingly we still fail to see any effect. In fact, it is only in Column 4, where the CBSA*time interaction dummies are included, that we find the negative and significant effect of negative equity on migration.

This discrepancy in results from using different sources of variation suggests that local house prices and other local economic conditions that affect migration are highly correlated across time, and even controlling for mean differences across CBSAs is insufficient to distinguish between the two influences[10]. Thus, if negative house price shocks are correlated with labor market shocks that disproportionately depress the migration propensities of renters, this would mask the effect of lock-in that owners are experiencing simultaneously. Given that the recession was particularly harmful to the disadvantaged socio-economic groups that also tend to be renters, it is perhaps not surprising that their migration opportunities declined in a way that aligned with the movement of local house prices[11]. Thus, only after systematically controlling for these labor market factors by including γ_{jt} can we isolate

[10] Abowd and Vilhuber (2012) find strong correlations between local house prices and labor market conditions as measured by the Quarterly Workforce Indicators, even when including CBSA and time fixed effects.

[11] Hoynes et al. (2012) find that the recession was most harmful to black, hispanic, young, and less educated workers.

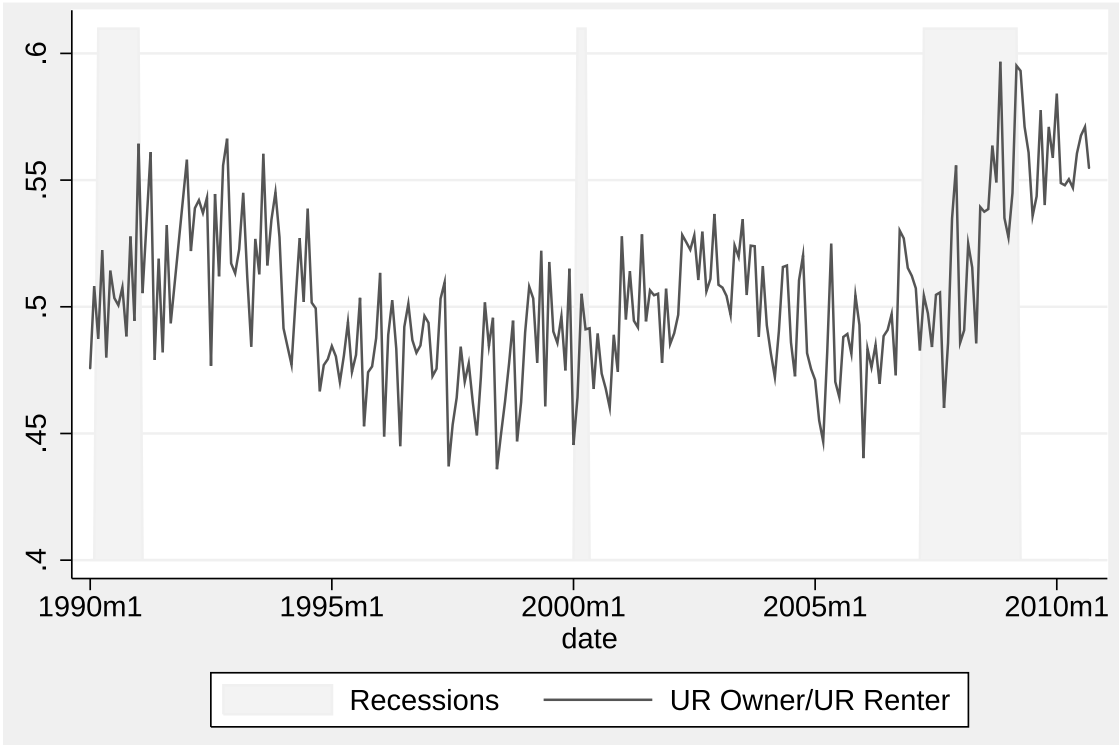
the impact on owners and see a feature of the data that is not otherwise possible.

2.6.2 Effect on the Unemployed

2.6.2.1 Regression Estimates

While the above results suggest that the lock-in effect may prevent many workers from making efficient job transitions, it could have an even more harmful effect on the unemployed, since their unemployment spells may potentially be prolonged due to their inability to move. Economists have long noted a link between homeownership and unemployment, as first suggested by Oswald (1997). This hypothesis that homeownership impedes the return to labor market equilibrium after shocks is supported by certain features of the data. An example, pictured in Figure 2.9, plots the ratio of the unemployment rate of homeowners to that of renters using monthly CPS data from 1990-2010, and it seen that this ratio increases during recessions and especially the most recent recession. Due to these reasons, commentators have expressed concerns that the effects of lock-in may disproportionately hit distressed areas that have experienced declines in both their housing and labor markets, contributing to unemployment persistence and a slower economic recovery.

Figure 2.9: Homeowner/Renter Unemployment Ratio: 1990-2010



To address this question of whether lock-in affects the mobility of the unemployed in particular, I repeat the within-CBSA-time test on the subset of individuals that the ACS classifies as being unemployed, which can be compared to results from subsamples of employed workers, and of those not in the labor force. The ACS contains an employment status variable, which codes an individual as unemployed if they report not having a job, but being available for and actively seeking work.

Results in Table 2.5 fail to show support for the lock-in phenomenon affecting the labor mobility of the unemployed. Column 1 shows that the negative equity interaction does not affect the out-migration rate of unemployed homeowners in a statistically significant way. Similarly, there is no significant impact on those who are labeled as not being in the labor force. The lack of significance may perhaps be attributable to the fact that the respective subsamples of unemployed and non-participants represent a fairly small fraction of the entire population, which is further exacerbated by the fact that the negative equity proxy is a binary variable that is also of relatively low frequency. The change in house price specification also shows no impact for the unemployed in Column 4, although it yields a marginally significant and positive coefficient for those not in the labor force in Column 5. Note that the observed variation in the log change in house price variable is greater than in the discrete negative equity proxy, perhaps affording greater precision.

Table 2.5: Results by Employment Status

	(1)	(2)	(3)	(4)	(5)	(6)
House Price measure	**Neg_equity**	**Neg_equity**	**Neg_equity**	**Δln(HPI)**	**Δln(HPI)**	**Δln(HPI)**
Subsample	**Unemp.**	**Not in L.F.**	**Emp.**	**Unemp.**	**Not in L.F.**	**Emp.**
	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)
HousePriceXOwn	**0.002**	**-0.000**	**-0.002*****	**0.002**	**0.004***	**0.005*****
	(0.003)	**(0.001)**	**(0.001)**	**(0.004)**	**(0.002)**	**(0.001)**
Own	-0.030*	-0.006	-0.006**	0.030*	-0.008	0.009***
	(0.010)	(0.008)	(0.002)	(0.012)	(0.008)	(0.02)
HousePrice	-0.001	-0.001	-0.000	-0.007*	-0.002	-0.001
	(0.002)	(0.001)	(0.000)	(0.003)	(0.002)	(0.001)
N	~300,000	~900,000	~6,100,000	~300,000	~900,000	~6,100,000

* p ≤ 0.05 , ** p ≤ 0.01, *** p ≤ 0.001

Notes: Sample consists of approximately 7.3 million ACS individuals i, either household head or spouse age 25-54, who are observed in one of 356 CBSAs j during a given quarter t in 2002q1-2012q4. All regressions include the following sets of controls: Ownership status X CBSA, CBSA X time period, Ownership status X time period. Individual controls for sex, age, race, education, marital status, presence of children, and residential tenure length are also included. Standard errors are clustered by CBSA.

Given the nature of the migration decision for the unemployed, it is arguably more appropriate to use a concept of migration that allows for a longer timing threshold. Since migrating job seekers may move before engaging in a job search, requiring that they be observed as employed within 1 quarter is perhaps too restrictive. Extending the threshold has the additional benefit of observing more migrators, which may alleviate some of the imprecision due to sparsely populated CBSA-time cells. Table 2.6 shows results for the unemployed using the 4-quarter threshold definition migration. We see that results are now significant at the 5% level in both specifications, with a coefficient of -0.6 in the negtive equity specification in Column 1, and 0.8 in the $\Delta ln(HPI)$ specification in Column 2. Given an average 4-qtr migration rate of around 10% for unemployed homeowners in the sample, this implies that the lock-in effect makes unemployed homeowners 8% less likely to migrate than average. This effect is smaller than the estimated 12% decrease for the entire sample. This diminished impact of lock-in perhaps reflects the increased incentives to relocate for the jobless, who may stand to gain more by moving than a migrator who is simply looking to switch jobs. Additionally, unemployed individuals face an increased risk of foreclosure as they may not have the necessary liquidity to continue to pay their mortgage. Both of these factors would serve to mitigate the lock-in effect, and help explain why the estimates are lower in magnitude for the unemployed.

Table 2.6: Unemployed Sample: 4-qtr Migration Threshold

House Price measure	(1) **neg_equity**	(2) **Δln(HPI)**
	β/(s.e.)	β/(s.e.)
$\Delta HousePriceXown_{ijt}$	**-0.006***	**0.008***
	(0.003)	**(0.004)**
Own_{ijt}	-0.014	-0.017
	(0.014)	(0.014)
$\Delta HousePrice_{ijt}$	0.003	-0.009**
	(0.002)	(0.003)
N	˜300,000	˜300,000

* p ≤ 0.05, ** p ≤ 0.01, *** p ≤ 0.001

Notes: Sample consists of approximately 7.3 million ACS individuals i, either household head or spouse age 25-54, who are observed in one of 356 CBSAs j during a given quarter t in 2002q1-2012q4. All regressions include the following sets of controls: Ownership status X CBSA, CBSA X time period, Ownership status X time period. Individual controls for sex, age, race, education, marital status, presence of children, and residential tenure length are also included. Standard errors are clustered by CBSA.

2.6.2.2 Impact on the Unemployment Rate

Using this point estimate from the unemployed sample in the 4-quarter migration specification, we can conduct a back-of-the-envelope calculation to estimate the impact of housing lock-in on the national unemployment rate. In the context of the canonical search and matching model framework developed by Diamond, Mortensen and Pissarides, Shimer (2007) states that the steady state unemployment rate can be expressed roughly as:

$$u_t \simeq \frac{s_t}{s_t + f_t} \tag{2.5}$$

where f is the job finding rate for the unemployed, and s is the rate of separation to unemployment for the employed. Therefore, assuming a constant separation rate s, the impact of housing lock-in on the equilibrium unemployment rate is a function of the difference between f and f_c, the counterfactual finding rate that would prevail in the absence of the lock-in friction:

$$\Delta u = \frac{s}{s+f} - \frac{s}{s+f_c} \tag{2.6}$$

In my data, the probability of an unemployed worker becoming re-employed within a year is 64%, with 52% finding a job within the same CBSA and 12% in a different CBSA. This is roughly consistent with estimates of the annual finding rate in the CPS as calculated by Cajner and Ratner (2014). Therefore, the estimated 8% reduction in the migration propensity for unemployed homeowners with negative equity implies a reduction in the annual job finding rate of about 1 percentage point for this group (assuming no difference in the finding rate in their original CBSA). However, since underwater homeowners only represent about one quarter of the unemployed in my sample at the depths of the market in 2010, this reduction in the job finding rate translates to only about a 0.25 percentage point reduction in the finding rate on an annual basis. The flow from employment to unemployment cannot be directly observed in my data set, but estimates from the CPS for the annual separation rate during the recession period of 2007-2009 is about 7%. Entering these numbers into the above equation results in an increase of the unemployment rate of less than a tenth of a percentage point. Therefore, even if lock-in does inhibit some unemployed individuals from relocating, it is not of a significant magnitude to affect the unemployment rate in a substantial way.

2.6.3 Robustness Analysis

2.6.3.1 Alternate Dependent Variables

For the sake of assessing the robustness of the above results, we will first explore using alternate definitions of the migration outcome. Table 2.7 shows results using the 2-quarter as well as the 4-qtr threshhold of migration, which allow for a longer job search process, intervening spell of unobserved unemployment, or employment at a job not covered by the LEHD system.

Table 2.7: Alternate Migration Thresholds

	2-qtr Threshold		**4-qtr Threshold**	
House Price measure	**(1)** **Neg_equity**	**(2)** **Δln(HPI)**	**(3)** **Neg_equity**	**(4)** **Δln(HPI)**
	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)
HousePriceXOwn	**-0.003*****	**0.006*****	**-0.004*****	**0.008*****
	(0.001)	**(0.001)**	**(0.001)**	**(0.002)**
Own	-0.008**	-0.010***	-0.008**	-0.011***
	(0.002)**	(0.002)	(0.003)	(0.003)
HousePrice	0.001	-0.002*	0.001	-0.003*
	(0.001)	(0.001)	(0.001)	(0.001)
N	˜7,300,000	˜7,300,000	˜7,300,000	˜7,300,000

* p $\leq$ 0.05, ** p $\leq$ 0.01, *** p $\leq$ 0.001

Notes: Sample consists of approximately 7.3 million ACS individuals i, either household head or spouse age 25-54, who are observed in one of 356 CBSAs j during a given quarter t in 2002q1-2012q4. All regressions include the following sets of controls: Ownership status X CBSA, CBSA X time period, Ownership status X time period. Individual controls for sex, age, race, education, marital status, presence of children, and residential tenure length are also included. Standard errors are clustered by CBSA.

The coefficient on the key interaction term for the 2-qtr migration threshold in column 1 is -0.3, while the corresponding coefficient in the 4-qtr specification is -0.4. Compared to the mean 2-quarter and 4-quarter migration rates for homeowners of 2.6 and 3.1 percent respectively, the implied changes in migration propensity due to negative equity are reduced by about 10-15% on average, similar to the previously shown estimates based on 1-quarter migration. Coefficients in the specifications using the log change of house prices are positive and significant, and also have similar implied magnitudes to previous results.

2.6.3.2 Differences in LEHD coverage

Because my measure of migration is based on the propensity of observing ACS individuals in the LEHD database, it is necessary to consider whether differences in workers' *a priori* probabilities of being observed could be influencing the results. First, due to the varying geographic coverage of the LEHD universe over time, we explore whether the results would differ if the baseline regression were run instead on a consistent sample of states. Using the measure of migration based on the 45 states present in the LEHD system as of 2002:Q1, as discussed above, Table 2.8 shows the results from such an analysis.

Table 2.8: Consistent 45-State Sample

House Price measure	(1) neg_equity	(2) Δln(HPI)
	β/(s.e.)	β/(s.e.)
$\Delta HousePriceXown_{ijt}$	**-0.002****	**0.005*****
	(0.001)	**(0.001)**
Own_{ijt}	-0.004	-0.006**
	(0.002)	(0.002)
$\Delta HousePrice_{ijt}$	0.000	-0.002*
	(0.001)	(0.001)
N	˜6,800,000	˜6,800,000

* p ≤ 0.05, ** p ≤ 0.01, *** p ≤ 0.001

Notes: Sample consists of approximately 6.8 million ACS individuals i, either household head or spouse age 25-54, who are observed in CBSA j during a given quarter t in 2002q1-2012q4. Sample is limited to 45 states that are present in the LEHD system as of 2002q1, and definition of migration is altered to include only those states as destinations. All regressions include the following sets of controls: Ownership status X CBSA, CBSA X time period, Ownership status X time period. Individual controls for sex, age, race, education, marital status, presence of children, and residential tenure length are also included. Standard errors are clustered by CBSA.

The point estimates of -.002 in the negative equity specification and .005 in the percentage change in house prices specification match those from using the complete sample, although the significance declines slightly in the negative equity specification, likely due to the lower number of observations. The overall similarity of these results, however, alleviates concerns that any bias is arising from the varying geographic coverage in the baseline sample.

A second coverage issue arises from the fact that certain types of individuals may be more or less likely to be accurately captured in the data. Not only are there classes of workers that are absent from the LEHD database, such as the self-employed, but there may also be unobserved characteristics that lead to a greater probability of being mismeasured. Previous research has shown that survey data do not always correspond to the administrative data in the LEHD system, and the reasons for such discrepancies are unclear (Abraham et al., 2009). Thus, combining survey and administrative data could create bias if the geographic and temporal distribution of mismeasured workers varies in such a way that is correlated with housing-related factors. To address this, I run the analysis on a subsample of individuals that are verified as having an LEHD job in the same reference quarter and CBSA of their ACS response, and who thus who demonstrate a tendency to be covered accurately by the LEHD universe. The fraction of such confirmed jobs in the sample is approximately 70%. Table 2.9 reports results from a subsample of "confirmed" jobs, compared to a subsample of those whose employment status in time t is "unconfirmed". In fact, we see that the lock-in effect is more pronounced in the confirmed cases than the unconfirmed cases, with coefficients on the negative equity interaction of -.003 compared to -.001, both significant at the 1% level. The same pattern is seen in results from the log change in house price specification. Thus while we have no specific reason to discount the unconfirmed cases, it is encouraging that the baseline results are being driven by the data points that display consistency between the two underlying sources.

Table 2.9: Sample by Whether Reference Job is LEHD-Confirmed

	(1)	(2)	(3)	(4)
House Price measure	**Neg_equity**	**Neg_equity**	**Δln(HPI)**	**Δln(HPI)**
Subsample	**Confirmed**	**Unconfirmed**	**Confirmed**	**Unconfirmed**
	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)
HousePriceXOwn	**-0.003****	**-.001****	**0.010*****	**0.002*****
	(0.001)	**(0.000)**	**(0.002)**	**(0.001)**
Own	-.017**	-0.003	-0.020***	-0.001
	(0.005)	(0.002)	(0.005)	(0.0031)
HousePrice	-0.000	0.001	-0.002	-0.004
	(0.001)	(0.000)	(0.002)	(0.002)
N	˜5,200,000	˜2,100,000	˜5,200,000	˜2,100,000

* p $\leq$ 0.05, ** p $\leq$ 0.01, *** p $\leq$ 0.001

Notes: Sample consists of approximately 7.3 million ACS individuals i, either household head or spouse age 25-54, who are observed in one of 356 CBSAs j during a given quarter t in 2002q1-2012q4. All regressions include the following sets of controls: Ownership status X CBSA, CBSA X time period, Ownership status X time period (results omitted for clarity). Individual controls for sex, age, race, education, marital status, presence of children, and residential tenure length are also included. Standard errors are clustered by CBSA

2.6.3.3 Demographic Heterogeneity

Given the design of the empirical framework, a key threat to identification would arise if house prices were proxying for something that has heterogeneous effects on the owner and renter groups. In particular, we may be worried that such unobserved factors will differentially influence people with different demographic characteristics. Given the very different compositions of the owner and renter pools, such an unobserved influence may be driving the different sensitivities to house prices of the two groups. Differences in their mean migration propensities by CBSA across time, as well as by time period nationally, are already accounted for with the existing controls. Thus this bias would only occur if the within-CBSA house price variation is correlated with other factors that may be affecting the migration propensities of certain groups of people across time.

In order to address these concerns about heterogeneous effects, Table 2.10 shows the results from re-estimating the baseline regression model on a variety of demographic subsamples. These specifications therefore represent full interactions between the homeownership and house price variables and the given demographic traits that may capture this heterogeneity. If the results remain fairly consistent across subsamples, this provides more reassurance that heterogeneous effects are not driving the baseline results.

Table 2.10: Coefficients for Demographic Subsamples

1) Sex	**Male**	**Female**		
Neg_Equity	-.003***	-.001		
Δln(HPI)	.006***	.005***		
2) Age	**25-34**	**35-44**	**45-54**	
Neg_Equity	-.002*	-.002*	-.001	
Δln(HPI)	.008***	.003**	.002	
3) Race	**White**	**Non-white**		
Neg_Equity	-.002***	-.002*		
Δln(HPI)	.006***	.005***		
4) Earnings	**1st Quartile**	**2nd Quartile**	**3rd Quartile**	**4th Quartile**
Neg_Equity	.001	-.001	-.002*	-.002**
Δln(HPI)	.005*	.004*	.005**	.004***
5) Marital status	**Married**	**Unmarried**		
Neg_Equity	-.003**	-.002***		
Δln(HPI)	.005***	.006***		
6) Children present	**Children**	**No children**		
Neg_Equity	-.001*	-.003***		
Δln(HPI)	.003**	.007***		
7) Education	**< Highschool**	**HS Diploma**	**Some College**	**College+**
Neg_Equity	.001	-.001	-.002*	-.002*
Δln(HPI)	-.002	.003	.006***	.006***

* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$

Notes: Sample consists of approximately 7.3 million ACS individuals i, either household head or spouse age 25-54, who are observed in one of 356 CBSAs j during a given quarter t in 2002q1-2012q4. All regressions include the following sets of controls: Ownership status X CBSA, CBSA X time period, Ownership status X time period. Individual controls for sex, age, race, education, marital status, presence of children, and residential tenure length are also included. Standard errors are clustered by CBSA.

The first panel of Table 2.10 estimates the model for males and females separately. The signs of the coefficients on the negative equity interaction are negative for both groups in columns 1 and 2, although the estimate is somewhat lower and statistically insignificant for females. However, the results using the log change in house price in columns 3 and 4 show strongly significant results for both males and females, and of a similar magnitude.

Panel 2 shows results from the three age categories, 21-34, 35-44, and 45-54. The results from the negative equity specification yield coefficients on the interaction term that are similar across groups and consistent with the baseline results, although the estimate for the 45-54 age category is not significant. The log change in house price specifications all yield positive coefficients, although of a somewhat higher magnitude for the youngest age group, and statistically insignificant for the oldest age category. Note that by splitting the sample into three or more subsamples, the lower number of observations does not afford the same precision as by using the entire sample, so more variability in the estimates is to be expected[12].

Estimates for whites and non-whites are shown in Panel 3 of the table. Here we see virtually identical point estimates across both specifications, although the negative equity specification yields slightly less significant estimates for the non-white group.

The results for the four earnings quartiles, shown in panel 4, reveal less precisely estimated coefficients due to the relatively small sample sizes, although all

[12]For complete regression results for these demographic subsamples, including standard errors, additional coefficients, and subsample sizes, please see Appendix A.

but 2 of the 8 specifications are at least significant at the 5% level. In Column 1 we see for the first time a coefficient of a sign contrary to the baseline results, but it is only very slightly positive and not statistically significant.

Panel 5 shows the results by marital status. The coefficients are very similar between the two categories, and comparable to baseline estimates. Some differences appear in the results based on the presence of children in the household, seen in Panel 6, as the lock-in effect appears to be stronger for those without children.

Finally, Panel 7 divides the sample by education level: less than high school, high school graduates, some college, and college graduates. Results are quite consistent across groups, although the coefficient is slightly positive and insignificant for the "less than high school" group in the negative equity specification, although note that this category contains the fewest observations. Across both specifications, the effect appears to be strongest for the more highly educated workers.

Across this wide set of demographic subsamples, the signs of the coefficients of interest have been of the same sign as in the baseline results in all but two cases. In addition, the magnitudes have generally been close to the baseline estimates of -0.002 in the negative equity specifications and 0.005 in the log change of house price specifications. Therefore taken as a whole, given the absence of widespread anomalies or outliers, these results show no particular cause for concern that heterogeneous effects on demographic groups may be confounding the baseline estimates.

2.6.3.4 Geographic Subsamples

Next we extend the analysis to additional subsamples based on geography, both in order to test for robustness as well as to explore some aspects of the lock-in phenomenon that may differ across areas for various reasons. First, we will test whether the strength of the lock-in effect differs across regions of the country which experienced the fall of house prices differently. Next, we explore whether the impact varied depending on whether individuals lived in states that prevented lenders from suing those who foreclosed for outstanding debt, a protection called "non-recourse" status. Finally, we will see if the lock-in effect exhibits differences in states with high rates of negative equity (and foreclosure) compared to those with lower negative equity rates.

While house prices declined across the country, some areas experienced the crash more strongly than others. While many notable examples occured in the cases of certain cities and states, such as Las Vegas and Arizona, these areas are too small to study with this within-CBSA methodology. However, there was significant house price variation across Census regions as well, and Table 2.11 shows the results from running the analysis separately on the Northeast, Midwest, South, and West regions.

Table 2.11: Results by Census Region Group

Subsample House Price measure	(1) Northeast Neg_equity	(2) Midwest Neg_equity	(3) South Neg_equity	(4) West Neg_equity	(5) Northeast Δln(HPI)	(6) Midwest Δln(HPI)	(7) South Δln(HPI)	(8) West Δln(HPI)
	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)
HousePriceXOwn	-0.003***	-0.001	-0.002*	-0.003**	0.004*	0.008*	0.005***	0.006**
	(0.0001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)	(0.001)	(0.002)
Own	-0.010**	-0.016***	-0.109***	-0.029***	-0.011**	-0.017***	-0.111***	-0.030***
	(0.003)	(0.001)	(0.005)	(0.002)	(0.004)	(0.001)	(0.005)	(0.003)
HousePrice	0.001	0.000	0.000	0.001	-0.002	-0.004	-0.002	-0.003*
	(0.001)	(0.001)	(0.000)	(0.001)	(0.0001)	(0.002)	(0.001)	(0.001)
N	~1,500,000	~1,800,000	~2,200,000	~1,800,000	~1,500,000	~1,800,000	~2,200,000	~1,800,000

* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$

Notes: Sample consists of approximately 7.3 million ACS individuals i, either household head or spouse age 25-54, who are observed in one of 356 CBSAs j during a given quarter t in 2002q1-2012q4. All regressions include the following sets of controls: Ownership status X CBSA, CBSA X time period, Ownership status X time period. Individual controls for sex, age, race, education, marital status, presence of children, and residential tenure length are also included. Standard errors are clustered by CBSA.

In the negative equity specifications in Columns 1-4, the magnitude of the key coefficient is highest for the Northeast and West, and lowest for the Midwest. This may suggest that the incentives to move away from the Rust-belt cities that were hit hardest by the recession, such as Detroit, overwhelmed the effects of lock-in. However, the log change in house price specifications in Columns 5-8 reveal the opposite pattern with the Midwest actually displaying the strongest lock-in effects, albeit somewhat imprecisely estimated. This implies that there may be non-linearities in the effects of changes in home equity.

Next we turn to a comparison of recourse vs. non-recourse states. As mentioned above, a non-recourse state is one in which the laws prevent mortgage lenders from suing a foreclosed homeowner for the difference between the proceeds of the foreclosure sale and the outstanding debt. Conversely, in recourse states lenders are allowed to seek redress in court, known as "deficiency judgements"[13]. In practice, however, such lawsuits have been rare, so it is unclear whether living in a recourse state substantially affects the incentives of underwater homeowners seeking to relocate. Table 2.12 shows that the coefficient for recourse states is indeed somewhat higher in magnitude in the negative equity specification, suggesting that recourse laws may be exacerbating the lock-in effect. However, the coefficients are identical in the log change of house price specification, although note that the recourse laws do not affect those with positive equity and therefore this specification may be less relevant than the specification that more closely identifies the at-risk group.

[13]The list of states with non-recourse laws are: Alaska, Arizona, California, Connecticut, Idaho, Minnesota, North Carolina, North Dakota, Oregon, Texas, Utah and Washington. The remainder are recourse states.

Table 2.12: Results by State Recourse Status

	(1)	(2)	(3)	(4)
House Price measure	**Neg_equity**	**Neg_equity**	**Δln(HPI)**	**Δln(HPI)**
Subsample	**Recourse**	**Non-Recourse**	**Recourse**	**Non-Recourse**
	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)
HousePriceXOwn	**-0.003*****	**-0.001***	**0.005*****	**0.005*****
	(0.001)	**(0.001)**	**(0.002)**	**(0.001)**
Own	-0.002	-0.008*	-.004	-0.009**
	(0.003)	(0.0013)	(0.003)	(0.003)
HousePrice	0.001	-0.000	-0.002*	-0.003**
	(0.001)	(0.0001)	(0.001)	(0.001)
N	˜2,400,000	˜4,900,000	˜2,400,000	˜4,900,000

* p ≤ 0.05, ** p ≤ 0.01, *** p ≤ 0.001

Notes: Sample consists of approximately 7.3 million ACS individuals i, either household head or spouse age 25-54, who are observed in one of 356 CBSAs j during a given quarter t in 2002q1-2012q4. All regressions include the following sets of controls: Ownership status X CBSA, CBSA X time period, Ownership status X time period. Individual controls for sex, age, race, education, marital status, presence of children, and residential tenure length are also included. Standard errors are clustered by CBSA.

Finally, Table 2.13 divides the sample by state-level negative equity rates. Specifically, using data from the Core Logic Negative Equity Report from 2010:Q4, each state is placed in "low", "medium" and "high" groups, where the cutoff levels for low and high are set at 30% and 15% respectively[14]. Again, the *a priori* effect of living in a high negative equity state compared to a low negative equity state is not clear. On one hand we might expect that since the high negative equity states are the ones hit hardest by the housing crisis, the lock-in effect is likely to be stronger. On the other hand, because high negative equity rates are accompanied by high rates of foreclosure, and because out-migrating from these most distressed areas may yield the largest benefit, it may be that the lock-in effect is in fact weaker in those areas. The estimates, however, do not show any substantial differences across

[14]High Group: Nevada, Arizona, Florida, Michigan and California. Medium Group: Georgia, Virginia, Maryland, Ohio, Idaho, Colorado, New Hampshire, Illinois, Utah, Rhode Island, District of Columbia, Massachusetts, Minnesota and New Jersey. Low Group: All other states.

these groups of states, either suggesting that there is no differential effect, or that the factors described above roughly offset one another.

Table 2.13: Results by State-Level Negative Equity Rate

	(1)	(2)	(3)	(4)	(5)	(6)
House Price measure	**Neg_equity**	**Neg_equity**	**Neg_equity**	**Δln(HPI)**	**Δln(HPI)**	**Δln(HPI)**
Subsample	**High**	**Medium**	**Low**	**High**	**Medium**	**Low**
	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)
HousePriceXOwn	**-0.002**	**-0.001**	**-0.002****	**0.005****	**0.006*****	**0.006***
	(0.001)	**(0.001)**	**(0.001)**	**(0.001)**	**(0.001)**	**(0.003)**
Own	-0.032	-0.004	-0.001	-0.033***	-0.007*	-0.006
	(0.006)	(0.002)	(0.002)	(0.006)	(0.003)	(0.003)
HousePrice	-0.000	0.000**	0.002	-0.002	-0.003*	-0.003
	(0.001)	(0.001)	(0.004)	(0.001)	(0.001)	(0.002)
N	˜1,900,000	˜2,100,000	˜3,300,000	˜1,900,000	˜2,100,000	˜3,300,000

* p ≤ 0.05, ** p ≤ 0.01, *** p ≤ 0.001

Notes: Sample consists of approximately 7.3 million ACS individuals i, either household head or spouse age 25-54, who are observed in one of 356 CBSAs j during a given quarter t in 2002q1-2012q4. All regressions include the following sets of controls: Ownership status X CBSA, CBSA X time period, Ownership status X time period. Individual controls for sex, age, race, education, marital status, presence of children, and residential tenure length are also included. Standard errors are clustered by CBSA.

2.7 Conclusion

This study has employed a novel method of measuring the impact of house price changes on the out-migration propensity of homeowners, by comparing the effects on owners and renters in the same location and time period. Results from the within-CBSA and time analysis reveal that the migration propensity of homeowners is strongly affected by changes in house prices. Specifically, homeowners who have experienced nominal declines in the value of their homes since moving in have migration rates that are 0.2 percentage points lower than comparable homeowners who have not experienced such a decline. This represents a figure that is 12% lower than the sample average for homeowners, and accounts for approximately 20% of the decrease in the overall decline in the mobility of homeowners from 2006-2010. The total percent change in house prices also appears to significantly affect migration propensity, suggesting that changes in equity matter for mobility on the positive as well as negative ends of the spectrum.

The effect of lock-in is less apparent for the unemployed, however, which in some ways may be the group for whom hindered mobility is the most harmful. This could partially be due to the small size of the sample of unemployed, which limits the precision of the estimates. It is also possible that the unemployed are less influenced by lock-in because of higher rates of foreclosure and increased incentives to relocate for work. However, further study is warranted to confirm that lock-in is not representing a significant friction for this group of workers. For example, it could be most worthwhile to investigate particularly distressed regions and industries, where the susceptibility of workers to lock-in would be particularly detrimental to economic recovery.

Nevertheless, taken as a whole, these results suggest that lock-in has a significant effect on national worker flows, and hence on the economy-wide reallocation of labor. The decline in labor flows in general has been a widely noted phenomenon as of late, and researchers have posited many potential consequences for labor market search, wage increases, and even output and productivity growth (Davis and Haltiwanger, 2014; Hyatt and Spletzer, 2014; Molloy and Wozniak, 2011). Investigating the causes of this decline will surely continue to be a much researched topic, and this study provides an example of how combining different sources of data can provide a valuable tool for exploring these complex labor market interactions.

Chapter 3: Unemployment Duration and Geographic Mobility: Do Movers Fare Better than Stayers?

3.1 Introduction

The ability to relocate to a different labor market has long been viewed as an important means by which the unemployed can improve their job search and escape unemployment more quickly. This issue was highlighted during the Great Recession, which was accompanied by a dramatic and well-documented decline in internal migration. While much of this decline was likely due to the lack of employment opportunities in general, researchers have argued that other factors such as a declining housing market further deterred unemployed workers from leaving the most distressed labor markets and thus prolonged the economic recovery. Implicit in these discussions, however, is the assumption that workers who move are better off for having done so, yet in fact there is scant empirical evidence that this is the case. While many models have focused on how migration helps effectuate equilibrium in unemployment rates and wages across regions, fewer studies have investigated whether the individual outcomes of the actual migrators are better than if they had remained in their original labor market. Thus, the literature is largely unsettled on whether migration by the unemployed is "micro-efficient", in the sense that every individual actor increases their utility, as discussed by Herzog et. al. (1993).

The migration literature has generally focused on measuring the influence of various factors in the migration decision, with DaVanzo (1978) being the first to focus on unemployment as a key determinant, providing evidence that the unemployed move at much higher rates than the average worker. Macroeconomic theories developed describing this phenomenon as the shifting of resources away from the less productive, high unemployment areas, to the more productive, low unemployment areas (Greenwood, 1975; Molho, 1986). Structural models of migration and unemployment followed, such as Harris and Todaro (1970), Sato (2004), Rogerson and McKinnon (2005) and Zenou (2009), describing how the interaction of unemployment and wages leads to equilibrium locational placement of workers.

Less attention has been paid, however, to the actual outcomes of the individual workers who move, especially in comparison to how they could have expected to fare by remaining in the same location. Finding evidence of the individual benefits of moving is important, since it is primarily through this incentive mechanism that equilibrium in the labor market can be achieved. Kennan and Walker (2011) employ a structural estimation using the NLSY, and show that workers' migration decisions are consistent with the maximization of lifetime wages, but the model does not focus on the short-term outcomes that may motivate the unemployed in particular. Those studies that do concentrate on the jobless population have generally failed to show much evidence of improved outcomes from relocating, with some even finding migration to have harmful effects on future outcomes. Both Shumway (1993) and Bailey (1994), using the SIPP and NLSY respectively, find that workers who migrate have somewhat longer spells of unemployment on average, although these studies are unable to rigorously address the endogenous selection into migration. Pekkala and

Tervo (2002) study unemployed workers in Finland and find that the employment rate of movers was much lower than that of stayers after the issue of self-selection into migration was controlled for. This study addresses the endogenous selection issue with an instrumental variables method using regional house prices as instruments for migration, a strategy that I will follow here.

In this study, I use a large sample of unemployed workers in the American Community Survey, combined with information about their employment outcomes from the Longitudinal Employer-Household Dynamics (LEHD) database, in order to measure the difference in re-employment propensities between those who move to another metropolitan area for a new job, and those who become employed in the same location as originally observed. To address the biases associated with selection into migration, I use information on local house price changes since each individual moved into his or her home, which should influence their migration decision but not be correlated with employability in general. Results from two-stage instrumental variables analyses show that those who move for new jobs become re-employed within a shorter time frame, a feature of the data that is not evident in more simple comparisons. These results imply that those who select into migration are relatively disadvantaged in terms of their ability to find work compared to those who stay, and stand to gain the most by relocating.

The discussion proceeds as follows. Section 2 describes the data and the method for measuring labor migration and unemployment spells. Section 3 discusses the empirical strategy, and motivates the use of house prices as an instrumental variable. Section 4 reports the results, which estimate the impact of migrating for a new job on the time to re-employment. Section 4 concludes.

3.2 Measuring Migration and Time to Re-employment

An unemployed individual residing in a particular location faces a decision whether to remain in the same labor market to find new employment, or else to relocate to another labor market. Analyzing these outcomes thus requires longitudinal data that allows us to observe an unemployed worker at time t in their original location j, and follow the individual until they find subsequent employment at time $t+k$, in either location j or another location $-j$.

To this end, I employ the combined ACS-LEHD dataset described in detail in Chapter 2. The large size of the sample is advantageous for this analysis, because the data sources more typically used to study migration, such as the NLSY, CPS, and SIPP, do not allow enough observations of unemployed people to generate very precise estimates for this group. By merging in the employment histories of these same individuals from the LEHD database, I can measure labor mobility amongst the ACS respondents and exploit the variation that the large sample affords[1].

The sample comprises all respondents i (reference person or spouse) aged 25-54 and residing in one of 356 CBSAs during any quarter t of the period 2002-2012, who are labeled as unemployed according to ACS definitions, and who also have no observed earnings in the LEHD database during quarter t. Note that because the ACS data is monthly and the LEHD data is quarterly, the restriction that the individual have no LEHD earnings during the entire quarter is more restrictive than

[1] See Chapter 2 for further details on the construction of this combined ACS-LEHD data set

the requirement that the respondent be labeled as unemployed during the month of the ACS interview. Approximately two-thirds of the unemployed in the ACS meet this restriction of zero LEHD earnings in t. Note that LEHD coverage varied geographically over the sample period, as more states began to participate during the middle part of the decade.[2] Thus I do not include in the sample any individual who is observed in the ACS during a time when their state of residence is not part of the LEHD data.

As before, these individuals are matched to their employment histories in the LEHD database at all UI-covered jobs through 2014q3. To determine when an ACS respondent unemployed at time t eventually becomes re-employed, I search for the first new LEHD job to begin after the observed date of unemployment. The next job with positive earnings beginning in a quarter $t + k$ is determined to be the new job. In case of multiple new jobs starting in the same quarter, the one with the greatest earnings is selected (i.e. the "dominant" job). Since a new job in a different CBSA may theoretically be accompanied by another new job beginning in the original CBSA, the concept of migration is somewhat clouded in this case, but the focus on the dominant job is meant to capture the most economically meaningful employment outcome.

The location of the job is given by the LEHD-assigned geographical information for the employer establishment, according to QCEW sources, thus allowing us to determine whether the new job is located in the same CBSA as reference CBSA

[2] As mentioned in Chapter 2, 45 states are covered at the beginning of the sample in 2002, 48 states by 2004, and 50 states plus D.C. by 2010.

j, or else in a different CBSA $-j$. Moving to a job located in an adjacent CBSA that is considered to be part of the same "Combined Statistical Area", or else to a rural area in the same state as j, is not counted as a move. As described in the previous chapter, the LEHD database imputes a worker's workplace establishment in the case where their employer has multiple establishments. This introduces error in my measurement of migration, although note that the LEHD imputation system is largely based on the proximity of the worker's residence to the employer establishment. This mitigates concerns that migration will be spuriously observed due to the imputation process.

The time to re-employment is defined as the number of quarters k between the observed reference period of unemployment t in the ACS survey, and the first quarter of subsequent employment in the LEHD data $t+k$. Note that the true beginning of the unemployment spell is not observed, because the ACS survey does not indicate when the period of unemployment actually started. While this means that the length of unemployment spells will be underestimated, assuming that the ACS interview occurs at a random point during an individual's unemployment spell, there is little concern that certain observations will be systematically more mis-measured than others. Also, because I do not observe jobs that are not covered by the LEHD universe, the end of unemployment spells may also be measured with error. The LEHD data were available through 2014q3 at the time of this analysis, therefore spells not completed by this time are right-censored.

Finally, because the definition of migration is based on employment and not place of residence, only completed unemployment spells will be included in the analysis. Therefore, the estimated effects of moving on re-employment will be conditional on eventual re-employment. This leads to a somewhat different interpretation compared to a study based on residential mobility, which can determine if an individual migrated regardless of whether he or she found employment. Such an analyis is a more direct test of whether moving is a mechanism that leads to employment, in the sense that moving can be viewed as a treatment (once selection is controlled for). However, my analysis offers a test of whether those who move and find employment appear to be doing so in a way that is consistent with the incentives to shorten their unemployment spell. As there has heretofor been little evidence that moving is beneficial for the individual actors, even such a narrowly tailored question is worth exploring.

Table 3.1 displays summary measures of re-employment in my sample, both for those who become employed in the same CBSA j as their residence in the original ACS reference period t ("stayers"), as well as for those who become re-employed in a different CBSA $-j$ ("movers"). Of the approximately 224,000 individuals in the sample of unemployed, 68% are observed becoming subsequently employed at a new job before the end of the sample in 2014q3. Due to right-censoring, the probability of observing re-employment is lower for those who are observed in the later years of the reference period 2002-2012, but note that yearly controls will be included in the main analysis to account for these differences.

Approximately 51.3% of the sample is next observed working in the same CBSA j as their ACS residence, while another 16% are next employed in a different CBSA $-j$. Consistent with previous findings, the average length of time to re-

Table 3.1: Re-Employment Statistics for ACS Sample of Unemployed

Variable	Mover	Stayer	All Re-employed
Fraction re-employed within 1 qtr.	0.044	0.155	0.199
Fraction re-employed within 2 qtrs.	0.070	0.240	0.310
Fraction re-employed within 4 qtrs.	0.102	0.341	0.443
Fraction re-employed by end of sample	0.166	0.513	0.679
Mean Time to Re-employment (qtrs.)	5.236	4.709	4.837
N	˜224,000		

employment is slightly longer for movers than for stayers, with average times to re-employment of about 5.2 quarters to 4.7 quarters respectively. The re-employment rate within 1 quarter is about 20% overall. Conditional on eventual re-employment, about 30% of stayers find jobs within the 1-quarter time frame, compared to about 25% of movers. This somewhat higher re-employment success rate for stayers also holds true for the 2-quarter and 4-quarter thresholds. Note that because of the above restrictions this sample represents a set of the long-term unemployed, therefore the re-employment propensities are lower than typically seen in the unemployment literature. However, the rise of the long-term unemployed was one of the dominant features of the Great Recession, and it is reasonable to believe that this is a group for whom the decision to move is particularly relevant[3].

While the focus of the analysis is on the individual benefits from relocating for employment, we will begin by investigating whether the aggregate out-migration propensities in our data appear to be related to external factors like local labor market conditions. Many empirical studies have noted a positive relationship between local unemployment rates and out-migration propensities, such as Basker (2003),

[3] As noted by Lawrence Katz in testimony before the Joint Economic Committee of Congress, April 2010.

Nakosteen et. al. (2008) and Haurin and Haurin (1998). To confirm this feature of the data in my sample, Table 3.2 calculates the observed out-migration rate for four categories of CBSAs, grouped by their average unemployment rate in 2009 according to CBSA-level data from the Bureau of Labor Statistics [4].

Table 3.2: Out-Migration by CBSA-Level Unemployment Rate: 2009

Unemployment Rate	**Out-Migration Fraction**
Under 9%	.149
9%-10%	.176
10%-11%	.218
Over 11%	.224

The four bins are centered around the approximate national unemployment rate of 10%. We see that the out-migration rates monotonically increase across the unemployment rate categories, implying that the labor markets that were most distressed during the Great Recession also experienced higher rates of out-migration. Indeed, out-migration from areas with unemployment rates exceeding 11% was 1.5 times greater than from areas with unemployment under 9%. These findings are consistent with the empirical and theoretical litereature, which highlight the role of migration in restoring equilibrium between labor markets during economic downturns.

3.3 Estimating the Impact of Migration on Re-employment

The central question of this study is whether out-migration for new employment shortens the time to re-employment. Clearly, the length of one's unemploy-

[4] Local Area Unemployment Statistics (http://www.bls.gov/lau/).

ment spell is not the only factor that an individual takes into account when deciding whether to move, but the notion that the unemployed make their locational choice based on their perceived chances of re-employment is common in the literature (DaVanzo, 1978; Herzog and Schlottmann, 1981). Indeed, while locational choice may well be influenced by long-term earnings prospects, the ability to restore one's stream of earnings as quickly as possible is of primary importance to the unemployed, who presumably have little or no income during the interim. Additionally, a growing literature has found that unemployment duration has harmful effects on earnings in the long-term as well as the short-term, further underscoring the incentives for shortening unemployment (Gregory and Jukes, 2001; Knight and Li, 2006; Cooper, 2014).

Migration as a means of job search has typically been modeled as a two-stage process as originally proposed by Sjaastad (1962). In the first stage, the unemployed individual must weigh the expected net present discounted value of moving versus staying, and chooses whichever yields the highest utility. The binary migration outcome M_i is thus expressed as:

$$M_i = \alpha Z_i + \epsilon \tag{3.1}$$

where Z_i represents all the variables that may influence the migration decision, including individual and household characteristics, as well as local and national economic conditions. The second stage expresses the realization of the employment outcome E_i, which in our case is the occurence of re-employment within a certain time frame:

$$E_i = \beta M_i + \delta X_i + u_i \tag{3.2}$$

where X_i is another set of individual-level and economic variables, potentially overlapping with Z_i, that affect the probability of re-employment.

As other previous studies have pointed out, one cannot simply estimate this system of equations via ordinary least squares (OLS), with bias arising due to the endogenous selection into migration (Bailey, 1991; Goss and White, 1994). In other words, migration cannot be considered to be a randomly allocated "treatment", because individuals choose whether to move based on factors related to their expected employment prospects in the different locations. If some of these factors are unobservable, ignoring these would result in a correlation between u_i and M_i, creating a biased estimates of β (and potentially δ). This bias could operate in either direction. For example, if the individuals who migrate tend to be those who have strong unobservable skills that increase their employment probablity, then uncorrected estimates will be upwardly biased, making migration appear more beneficial for employment outcomes than it truly is. Conversely, if movers are the most disadvantaged workers who have the poorest re-employment prospects generally (conditional on observables), then such estimates will consequently be biased downward. (Nakosteen et al., 1980; Herzog et al., 1993).

Many empirical methods exist to correct for such selection bias, the most commonly used being based on instrumental variables in a linear probability model, as described in Angrist (2001). If an instrument can be found that is both independent of u_i and not directly related to E_i conditional on X_i, then β can be estimated consistently in the two-stage-least-squares (2SLS) framework. Similar to 2SLS are various treatment effect models, which adapt the well-known Heckman (1979) correction to handle endogeneity on the explanatory variable rather than on the dependent variable (Maddala, 1999). Recently the use of such models has been called into question, as their non-linear nature makes them particularly sensitive to the assumptions of normality and heterskedasticity (Deaton, 1997; Angrist, 2001). Control function-based models are also employed, such as the two-stage Probit regression, although this have been shown to be inconsistent in the case of binary endogenous regressors. Hazard models adapted for endogenous regressors also exist, but are seldom used due to their computational intensity. Thus, while the emphasis in this study will be on results from a two-stage least squares linear probability model, other specifications will be explored for robustness.

The instrument for migration that I employ is a measure of the change in local house prices that the individual has experienced since living in their home. Home prices have long been seen as an important factor in the migration decision, with many studies examining the relationship between house prices and flows of migrants at the aggregate level. This literature has highlighted the role of regional house prices as a marker of the relative cost of living between different cities, as well as a reflection of the value of amenities in those locations. (Graves, 1983, Evans, 1990; Gabriel et. al., 1992, Mueser and Graves, 1995; Bitter, 2008). Experiencing house price movements during one's tenure is therefore likely to have a strong impact on the decision to migrate for these reasons. Such a change in prices directly influences the affordability of remaining in one's home area compared to moving elsewhere. Additionally, these movements in local house prices reflect changes in the desirability of an individual's neighborhood, due to changing amenities and other environmental factors. Finally, house price movements affect the amount of equity that homeowners have in their homes, which has been shown to have an impact on one's ability to move, as explored in Chapter 2. Note that besides the equity effect, most of the influences of home prices on migration listed here are relevant to renters as well as home owners. In fact, some channels may be felt even more strongly by renters. For example, the affordability of an area is more directly tied to home prices for renters, whose rent is variable, compared to homeowners who likely have fixed monthly housing costs. For the sake of tractability, however, the house price variable will be assumed to have a common effect on both owners and renters.

This house price instrument is presumed to be valid because while house prices affect the propensity to move, as discussed above, they should be uncorrelated with employability in general once individual characteristics and other labor market related controls are included. This identification strategy is similar to that of Pekkala and Tervo (2002) in their study using adminstrative data on the Finnish workforce. Unlike in that study, however, I choose not to use homeownership itself as an instrument for migration, because the decision to purchase a home in a given area could be related to one's unobservable abilities to find employment in that area relative to elsewhere. House-price changes are more plausibly exogenous since they are outside of the individual's control.

To construct this instrument, I define a variable expressing the change in house prices that an individual has experienced since moving into their home, in the same way as described in Chapter 1. Specifically, I define the variable

$$\Delta ln(HPI_{it}) = ln(HPI_{z_j t}/HPI_{z_t^0}) \tag{3.3}$$

where t is the reference time period, t^0 is the time period when the individual moved into their home according to the ACS, and HPI represents Zillow's house price index for the individual's zip-code of residence z (within CBSA j) during the corresponding points in time [5]. Note that whereas in Chapter 2 this interaction term was the key independent variable of interest, in the current analysis it will be used as an instrumental variable for the propensity to migrate.

Using this instrument based on individual-level house price changes, I will estimate the causal impact of moving for new employment on the probablity of re-employment within a given time-frame. Note that this approach takes the initial unemployed status of the individuals as given, and does not consider how the individuals arrived at that state in the given place and time. Some have suggested, such as Vijverberg (1995), that this potentially creates bias if individuals select into unemployment in a way that is related to one's job prospects in a given location. Due to data limitations, I am unable to tell much about the circumstances of the individual's unemployment. However, because of our sample restriction that the individuals have no LEHD earnings during the calendar quarter, and because the

[5] For more information on the Zillow data and the construction of the change in house price variable, see Chapter 2.

ACS definition of unemployment requires that the individual be "actively seeking work", it is reasonable to assume that we observe a fairly homogeneous sample of the involuntary unemployed, rather than some sort of "strategic" unemployment.

3.3.1 Regression Model

We now turn to the regression specifications designed to estimate the impact of migration on the propensity to become re-employed within a given time frame. The outcome of interest is the achievement of new employment within a particular threshold. Thus, the dependent variable is denoted as *emp_within_1q*, an indicator for whether the respondent is re-employed within 1 quarter of the observed date of unemployment t in the ACS. I will also use re-employment within 2 quarters (*emp_within_2q*), 4 quarters (*emp_within_4q*), and the total number of quarters before the termination of the unemployment spell (*unemp_duration* (*qtrs.*)), as alternative left-hand-side variables. The independent variable of interest in all specifications is $Mover$, an indicator for whether the worker is observed becoming re-employed in a CBSA $-j$ other than the CBSA j where they were originally observed residing in the ACS.

Key controls will include personal and household characteristics that are known to influence the migration decision (Greenwood, 1997). For instance, age has been found to be a strong determinant of migration, with older workers less likely to be observed relocating (Sjastaad, 1962). Household factors such as marital status, presence of children in the household, and homeownership, have also been discovered to lower the probability of moving (Sandell, 1977; Mincer, 1978; Graves and Linneman, 1979). Human capital-related factors such as education are likewise important, with empirical evidence showing that higher-educated workers tend to be more mobile (Basker, 2003; Wozniak, 2010). The analysis will therefore include measures of age, sex, race, education, marital status, presence of children, and homeownership, all of which are provided on the ACS. No earnings variable will be included since by definition these individuals are not presently employed.

Another individual-level control that I include is the number of years that a person has lived in their current home, represented by a slate of yearly dummy variables. This will allow the effects of housing tenure length on migration to vary in a non-parametric fashion. This control is especially important since the length of tenure is one of the sources of variation in the house price variable, and thus it serves to prevent the influences of the house price variable from being conflated with the influences of housing tenure length.

Local labor market conditions and other economic factors will be accounted for through other right-hand-side controls. Yearly dummies will account for national shocks, such as recessions, which will naturally impact national migration propensities and re-employment rates simultaneously. Dummies for each of the 356 CBSAs are also included to control for the mean levels of out-migration and re-employment at the local level. As discussed in the previous chapter, the LEHD coverage expanded during the sample period, therefore the probability of observing an individual moving will consequently vary across time and geographic location. However, the year and CBSA dummies should help account for the fact that the average propensity to be detected in a new location will differ by year and CBSA. While the sample of unemployed is not large enough to accommodate the inclusion of CBSA X time controls, as was done when using the entire sample of the ACS workforce in the previous chapter, I will attempt to control for time-varying, location-specific shocks by including the unemployment rate in the individual's home CBSA j according to BLS's Local Area Unemployment statistics.

The outcome equation can thus be expressed as:

$$Emp\_Within\_k_{it} = \beta' Mover_{it} + \delta' X_{it} + ur_{jt} + \phi_j + \gamma_t + u_{it} \tag{3.4}$$

where X_{it} are the individual and household-level controls, ur_{jt} is the unemployment rate in the reference CBSA and time period, and ϕ_j and γ_t are the CBSA and year dummies respectively. This specification can be estimated via OLS, which will produce biased estimates as discussed above, but will serve as a baseline against which to compare our selection-adjusted results. To adress the endogeneity, I estimate the following first-stage equation for migration:

$$Mover_{it} = \alpha' Z_{it} + ur_{jt} + \phi_j + \gamma_t + \epsilon_{it} \tag{3.5}$$

where Z_{it} is the instrument set which includes all the variables in X_{it}, in addition to the excluded instrument $\Delta ln(HPI)$. In the two-stage least squares estimation, the predicted values from the first stage $\widehat{Mover}_{it}$ are substituted for $Mover_{it}$ in equation 3.4, which can then be solved to obtain unbiased estimates of β. Results are consistent provided that the identifying assumptions of $Cov(Mover, Z) \neq 0$, and $Cov(Z, \epsilon) = 0$, hold true. All standard errors are clustered by CBSA, in order to obtain estimates that are robust to broad patterns of error correlation within labor markets.

3.4 Results

For the analysis, the sample must be limited to completed unemployment spells, since I am only able to identify workers as either movers or stayers if and when they become re-employed. As mentioned above, this means that the analysis measures the impact of moving on the time to re-employment *conditional* on eventual re-employment. After removing individuals who never receive LEHD earnings subsequent to observed unemployment, the sample size is reduced by roughly one third to approximately 152,000. Summary statistics for the final sample of completed unemployment spells are shown in Table 3.3. We now see that conditional on re-employment, about one quarter of re-employed workers are observed employed in another CBSA, while three-quarters remain in their original home CBSA. Approximately 29% of individuals are re-emloyed within 1 quarter, 46% within 2 quarters, and 65% within 4 quarters.

Table 3.3: Regression Sample: Summary Stats

Variable	**Mean**	**Std. Dev.**
$Mover$	0.244	0.429
$Stayer$	0.756	0.429
$Age25-34$	0.298	0.458
$Age35-44$	0.337	0.473
$Age45-54$	0.365	0.481
$Female$	0.549	0.498
$Non-white$	0.287	0.452
$Less\_than\_Highschool$	0.129	0.335
$Highschool$	0.276	0.447
$SomeCollege$	0.347	0.476
$College+$	0.248	0.432
$Married$	0.621	0.485
$Children$	0.563	0.496
$Owner$	0.568	0.495
$\Delta HousePrice$	0.233	0.371
Emp_within_1q	0.293	0.455
Emp_within_2q	0.457	0.498
Emp_within_4q	0.652	0.476
$Unemp\_duration(qtrs.)$	4.837	5.388
Sample Size:˜152,000		

3.4.1 Results From Linear Models

Baseline results from the linear specifications are shown in Table 3.4. Column 1 displays results from the OLS specification. The coefficient on *mover* reveals a 1-quarter re-employment rate for those who migrate that is roughly 5 percentage points lower than for those who stay. This slightly negative impact of moving is consistent with previous findings in the literature, such as Shumway (1993), Bailey (1994), and Pekkala and Tervo (2002). However, the IV results in Column 2 reveal strongly *positive* and significant effects of moving on 1-quarter re-employment, with a coefficient of 0.585. These IV results stand in contrast to those of Pekkala and Tervo (2002) who employ a similar IV specification and find even stronger negative effects than in their OLS estimates.

Table 3.4: Regressions: New LEHD Employment Within 1 quarter

	1 : OLS $\boldsymbol{\beta}$/(s.e.)	**2 : 2SLS – IV** $\boldsymbol{\beta}$/(s.e.)
Mover	**-.049*****	**0.620*****
	(.004)	**(.191)**
Own	.010* **	.033***
	(0.003)	(.007)
$Age35-44$	-.020***	-.017***
	(.003)	(.004)
$Age45-54$	-.033***	-.029***
	(.004)	(.006)
Female	-.018***	.006
	(.003)	(.009)
Non-white	-.022***	-.017***
	(.004)	(.005)
Married	.012***	.012***
	(.003)	(.004)
Children	-.017***	-.005
	(.003)	(.005)
Highschool	.016***	.012**
	(.005)	(.005)
SomeCollege	.015**	.004
	(.005)	(.006)
College+	.024***	.001
	(.006)	(.009)
Unemployment Rate	-.006	-.010*
	(.003)	(.004)
$CBSA\,and\,Year FE$	yes	yes
N	~152,000	~152,000

* p $\leq$ 0.05, ** p $\leq$ 0.01, *** p $\leq$ 0.001

Notes: Sample consists of approximately 152,000 completed unemployment spells by ACS respondents i, age 25-54 and residing in one of 356 CBSAs j during 2002-2012. Column 1 reports estimates from an OLS model, and Column 2 shows results from a 2-step Instrumental Variables regression. All regressions include dummy variables for the reference CBSA and year of unemployment. Individual controls for sex, age, race, education, marital status, presence of children, and residential tenure length are also included. All standard errors are clustered by CBSA.

Note that aside from the *mover* variable, the coefficients on the other variables do not differ significantly between the OLS and 2SLS specifications, with virtually none reversing sign. Homeowners, younger workers, married people, and the higher educated have higher re-employment propensities, while minorities, females, and households with children have somewhat lower re-employment propensities. Higher unemployment rates in one's home location are also associated with lower re-employment propensities.

Table 3.5 contains results from the first stage regression, revealing a coefficient of -0.038 on the excluded house-price instrument $\Delta HousePrice$ that is strongly significant and lies within the unit interval. This expresses the influence of home prices that is common to both owners and renters, and implies that home price appreciation is generally associated with less out-migration. This suggests that the rising amenity values that are reflected in higher prices provide a strong incentive to staying put, a phenomenon that Pekkala and Tervo (2002) also observe in their data.

Table 3.5: First Stage IV Results

	$\boldsymbol{\alpha}$/(s.e.)
$\Delta ln(HPI)$	**-0.032***** **(0.004)**
Own	-0.031*** (0.003)
Age36-45	-0.003 (0.003)
Age46-55	-0.002 (0.005)
$Female$	-0.036*** (0.004)
Non-$white$	-0.009** (0.003)
$Married$	-0.001 (0.003)
$Children$	-0.018*** (0.002)
$Highschool$	0.006 (0.005)
$SomeCollege$	0.015** (0.005)
$College+$	0.032*** (0.006)
$UnemploymentRate$	.004* .002
$CBSA\,and\,Year\,FE$	yes
N	˜152,000

* p $\leq$ 0.05, ** p $\leq$ 0.01, *** p $\leq$ 0.001

R-sq.	Adjusted R-sq.	Partial R-sq.	Robust F(3,232)	Prob > F
0.0806	0.0784	0.0006	36.0622	0.0000

Notes: Table reports results from the first stage regression of the Z instruments on migration. All right-hand-side variables from the second stage regression reported in the previous table are included, in addition to the excluded instrument $\Delta ln(HPI)$.

Note also that homeownership itself, aside from the influence of prices, has a strong negative effect on out-migration. Recall that homeownership is an independent variable and not an excluded instrument because it is presumed to potentially be directly related to re-employability. Finally, we see that the F-statistic from the first-stage of the IV regressions is 72, alleviating concerns of weak-instrument bias according to the criteria put forth by Staiger and Stock (1997).

Other coefficients in the first-stage regression are also consistent with the previous literature on the determinants of migration. For example, we see that the older age groups are associated with lower mobility, compared to the excluded 25-34 group (although the results are not significant). Females, minorites, and households with children are also less likely to relocate. The education categories exhibit the strongest relationship to migration, however, with the coefficient on $College+$ suggesting an increased migration propensity of over 3 percentage points. This corresponds to the findings of Malamud and Wozniak (2012), which showed a large impact of post-secondary education on migration propensities.

Table 3.6 shows the results using alternate dependent variables, representing re-employment in 2 quarters, re-employment in 4 quarters, and the total length of the completed unemployment spell (in quarters). All three specifications reveal the same pattern, with OLS estimates suggesting that moving for employment results in poorer re-employment rates and longer unemployment durations, while IV results imply that moving has beneficial effects on outcomes. Only the coefficient in the Emp_within_2q specification is significant at the 1% level, however, with the other two coefficients significant at the 5% level.

Table 3.6: Regressions with Alternate Dependent Variables

Dependent Variable	**1: OLS – Coeff. on *Mover*** $\boldsymbol{\beta}$/(s.e.)	**2: IV – Coeff. on *Mover*** $\boldsymbol{\beta}$/(s.e.)
Emp_within_2q	-.062***	0.780***
	(.005)	(.183)
Emp_within_4q	-.063***	.567**
	(.006)	(.240)
Unemp_duration (qtrs.)	.763***	-12.081*
	(.094)	(5.985)
$CBSA and Year FE$	yes	yes
N	˜152,000	˜152,000

* p ≤ 0.05, ** p ≤ 0.01, *** p ≤ 0.001

Notes: Sample consists of approximately 152,000 completed unemployment spells by ACS respondents i, age 25-54 and residing in one of 356 CBSAs j during 2002-2012. Column 1 reports estimates from an OLS model, and Column 2 shows results from a 2-step Instrumental Variables regression. All regressions include dummy variables for the reference CBSA and year of unemployment. Individual controls for sex, age, race, education, marital status, presence of children, and residential tenure length are also included. All standard errors are clustered by CBSA.

While these findings are clear in suggesting that migrating for employment results in a sharp reduction in unemployment duration, there are reasons to suspect that these results may not be universally generalizable given the large magnitude of the coefficients. For instance, note that the IV estimates imply that a mover improves her 1-quarter employment probability by 62% and shortens her unemployment spell by 12 quarters. According to the "local average treatment effect" interpretation of IV results as discussed in Angrist and Imbens (1995), these estimates are identified off of individuals for whom the moving decision is determined at the margin by the change in their home price. Whether these people are representative of the general population is unknown. In addition, the fact that the sample is composed of the longer-term unemployed gives further reason to believe that these results may not apply to the unemployed population as a whole.

Note also that because we are considering the impact of moving conditional on re-employment, the estimates apply to a particular subset of unemployment spells and are therefore not directly measuring how moving affects re-employment in general. Further work using data on residential location, in combination with employment outcomes, would be needed to test the treatment efect of moving. Nevertheless, these IV results suggest that a subset of unemployed workers who migrate and obtain employment are doing so in a way that is consistent with the incentive to shorten their unemployment spell. The discrepancy with the OLS results suggests that such movers possess unobservable qualities that make them less well-suited to finding employment in their original location. This provides novel evidence that individuals with poor employment prospects in a given location are properly incentivized to go elsewhere for employment, in accordance with theory.

3.4.1.1 Comparison to Previous Findings

It is worth considering the large contrast between the results from the IV regressions that we see here and those of Pekkala and Tervo (2002), who employed a similar IV strategy using Finnish administrative data. As mentioned, their IV results showed even more strongly negative effects than in their corresponding OLS specifications. Contrary to my conclusion, the implication of their findings was that migrators are, in fact, the *best* suited to finding re-employment generally, and were actually extending their time to re-employment by moving.

Besides the obviously different samples in terms of populations and time frames, there are a couple notable differences between these two studies. First of all, Pekkala and Tervo (2002) use regional house prices as an instrument, while I construct an individual-level house price change based on sub-metropolitan level data. Secondly, the Finnish study is able to measure residential mobility, lending itself to to a somewhat different interpretation than my analysis, as discussed above. Finally, and perhaps most importantly, Pekkala and Tervo chose to use homeownership itself as an excluded instrument, whereas I do not. I make this decision based on the assumption that the home ownership decision could well be related to unobservable factors that influence employability, in particular, employability in one's home location relative to other areas. In other words, people may choose to buy a house based on their private information and preferences that make them especially well-suited to finding employment in that area. In this case, omitting homeownership from X and including it in Z would violate the exclusion restriction $Cov(Z, \epsilon) = 0$, because homeownership's direct impact on employability would then be captured in ϵ. As the IV estimates are identified off the marginal mover who relocates because they are a renter, comparing their employment outcomes to those of a similar homeowner who stayed in their home location leads to biased results, since the homeowner actually possessed an unobserved advantage in employability.

Table 3.7 explores whether this difference in the instrument set is the cause of the discrepancies between the results in our two analyses. Column 1 shows the already reported coefficients on *mover* from the 2SLS specifications for the various dependent variables, while Column 2 repeats the analysis with the inclusion of home ownership as an additional excluded instrument.

Table 3.7: IV Regressions with Alternate Instrument Sets

Dependent Variable	**1: Original Instruments** **Coeff. on *Mover*** $\boldsymbol{\beta}$/(s.e.)	**2: Add Owner Instrument** **Coeff. on *Mover*** $\boldsymbol{\beta}$/(s.e.)
Emp_within_1q	0.620***	-.058
	(.191)	(.096)
Emp_within_2q	0.780***	-.003
	(.183)	(.108)
Emp_within_4q	.567**	.003
	(.240)	(.117)
Unemp_duration (*qtrs.*)	-12.081*	-3.118
	(5.985)	(2.289)
$CBSA and Year FE$	yes	yes
N	˜152,000	˜152,000

* p ≤ 0.05, ** p ≤ 0.01, *** p ≤ 0.001

Notes: Sample consists of approximately 152,000 completed unemployment spells by ACS respondents i, age 25-54 and residing in one of 356 CBSAs j during 2002-2012. Column 1 reports estimates from an OLS model, and Column 2 shows results from a 2-step Instrumental Variables regression. All regressions include dummy variables for the reference CBSA and year of unemployment. Individual controls for sex, age, race, education, marital status, presence of children, and residential tenure length are also included. All standard errors are clustered by CBSA.

We see that the results from the two specifications are indeed markedly different. The inclusion of home ownership as an instrument pushes the point estimates of the $Mover$ coefficients back toward zero, reversing the sign in 3 out of the 4 cases. None of the coefficients are significant, however, and none are farther from zero than their corresponding OLS estimates. Thus, we still do not see evidence that the OLS estimates are upwardly biased, as Pekkala and Tervo (2002) found. Given the imprecision of the results, it is difficult to come to any definitive conclusions, but it appears likely that the decision whether to use homeownership as an instrument is partly contributing to the large gap between the results from the two analyses.

3.4.2 Results from Probit models

As robustness tests for the baseline results, I explore other, non-linear, specifications. First we explore a Probit model, which are designed to handle binary outcomes such as our re-employment variables. However, note that the 2-stage probit model that is most analogous to the setup of two-stage-least-squares is not technically appropriate for use with binary endogenous regressors (Wooldridge, 2010) . Thus, this example will serve illustrative purposes only.

The two-step Probit model can be applied to equations (4) and (5) above, by adding the following further assumption:

$$(u_{it}, \epsilon_{it}) \sim N(0, \Sigma) \tag{3.6}$$

where σ_{11} is normalized to equal 1 for purposes of identification. This can then be estimated via the two-step method described in Newey (1987).

Table 3.8 shows the results of a standard Probit regression on equation (4) in the left-hand column, and of the two-stage Probit with instrumental variables in the right-hand column. The unadjusted estimates in Column 1 show a negative impact of migration on re-employment, while the 2-stage IV Probit results show strong positive effects – mirroring the pattern seen in the results from the linear specifications. Using the 2-quarter or the 4-quarter thresholds yields similar qualitative results. (Note that time-to-remployment is not a binary outcome and thus cannot be estimated via Probit.)

Table 3.8: Probit Specification

Dependent Variable	**1: Probit** **Marg. Effect of *Mover*** $\boldsymbol{\beta}$/(s.e.)	**2: IV Probit** **Marg. Effect of *Mover*** $\boldsymbol{\beta}$/(s.e.)
Emp_within_1q	-.050***	1.43***
	(.004)	(.279)
Emp_within_2q	-.064***	1.508***
	(.005)	(.212)
Emp_within_4q	-.064***	1.252***
	(.006)	(.372)
$CBSA and Year FE$	yes	yes
N	˜152,000	˜152,000

* p $\leq$ 0.05, ** p $\leq$ 0.01, *** p $\leq$ 0.001

Notes: Sample consists of approximately 152,000 completed unemployment spells by ACS respondents i, age 25-54 and residing in one of 356 CBSAs j during 2002-2012. Column 1 reports estimates from a Probit model, and Column 2 shows results from a 2-step IV Probit. All regressions include dummy variables for the reference CBSA and year of unemployment. Individual controls for sex, age, race, education, marital status, presence of children, and residential tenure length are also included. All standard errors are clustered by CBSA.

3.4.3 Results from Hazard models

Unemployment spells are known to exhibit duration dependence and have thus been commonly studied in a hazard or survival framework since Meyer (1990, 1991). Such models systematically account for the baseline hazard, or the natural propensity to escape unemployment at a certain point in time, given the survival of the unemployment spell up to that point. Unfortunately, there are few hazard models built to handle endogenous regressors, and those that are available require restrictive assumptions and are computationally onerous.

To explore this avenue, however, I will follow the method created by Bjiwaard and Ridder (2005), a linear rank estimator based on a mixed proportional hazard model with a piecewise linear baseline hazard. This model accomodates the use of instruments for the endogenous covariate, although only one instrument is permitted. As a consequence, I choose to include only the $\Delta HousePrice$ variable and omit its interaction with home ownership. Estimation requires intensive numerical maximization, via Newton-Raphson and Powell methods, thus I run the model on a random sample of only 3,000 observations, and omit the year and CBSA fixed effects.

Table 3.9 shows the results from the hazard analysis where the dependent variable is emp_within_1qtr. To provide a baseline for comparison, Column 1 shows the results from a unadjusted hazard regression, with a parametric baseline hazard defined by the exponential distribution. Coefficients are translated to express marginal efects at the average values of covariates. The coefficent on *mover* in the exponential model is -.085, implying that migration reduces the propensity to escape unemployment. The results of the Bijwaard and Ridder estimator in Column 2, however, show that the sign of the *mover* coefficient has reversed, with a coefficient of .185 that is significant at the .1% level, once again expressing a positive influence of migration on re-employment. The signs on the coefficients of the other covariates do not change between the two specifications, and are largely consistent with the linear 2SLS results. One exception is the estimated impact of the local unemployment rate, shown here to be positively correlated with re-employment, but this could potentially be due to the omission of the year and CBSA controls which were present in the earlier tests.

Table 3.9: Hazard Specification

	1 : ExponentialHazardModel $\boldsymbol{\beta}$/(s.e.)	**2 : MixedProp.Hazard − IV** $\boldsymbol{\beta}$/(s.e.)
Mover	-.085*	.183***
	(.043)	(.004)
Homeowner	.032	.022***
	(0.006)	(.002)
*Age*35 − 44	.011	.017***
	(.046)	(.001)
*Age*45 − 54	-.058	-.046***
	(.048)	(.001)
Female	-.040	-.013***
	(.037)	(.001)
Non-white	.001	-.003***
	(.041)	(.001)
Married	-.048	-.025***
	(.041)	(.002)
Children	-.082*	-.062***
	(.041)	(.002)
Highschool	.092	.063***
	(.064)	(.002)
SomeCollege	.085	.055***
	(.061)	(.002)
College+	.123	.084***
	(.065)	(.002)
Unemployment Rate	.032***	.015***
	(.006)	(.000)
CBSA and Year FE	yes	yes
N	˜3,000	˜3,000

* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$

Notes: Sample consists of 3,000 completed unemployment spells by ACS respondents i, age 25-54 and residing in one of 356 CBSAs j during 2002-2012. Column 1 reports estimates from an exponential hazard model, and Column 2 shows results from a mixed proportional hazard IV model. All regressions include dummy variables for the reference CBSA and year of unemployment. Individual controls for sex, age, race, education, marital status, presence of children, and residential tenure length are also included. All standard errors are clustered by CBSA.

3.5 Conclusion

This study combines information from the ACS and the LEHD database to measure unemployment duration and out-migration in a sample of the unemployed. The focus of the analysis was to test how migrators fare in terms of their time to re-employment compared to how they would have done had they reamained in their original location. By controlling for selection into migration with an instrumental variable strategy based on local house price changes, I find that movers have a significantly *higher* probability of re-employment within a given time threshhold, and *shorter* unemployment spells. This pattern is seen in estimates from a 2SLS linear probability model, as well as in non-linear specifications such as Probit and hazard models. While these findings stand in contrast to some of the previous literature, we note that these results will naturally be sensitive to different choices in instruments. Here I argued that instruments based on individual house-price changes are a parsimonious and plausibly exogenous instrument for migration, but future work in finding other instruments or exogenous variation in migration would be valuable for assessing these issues further.

Ultimately, it should perhaps not be surprising that migration appears to improve the outcomes of the unemployed. Relocating is a costly endeavor in terms of money and time, which the jobless can often ill-afford. Leaving one's home area may also represent a loss of location-specific capital in the labor market. Therefore, in order to provide sufficient incentive to undertake a long-distance move, the benefit from doing so must be large enough to outweigh the potentially substantial costs.

It stands to reason that this is most likely to hold true for those who are struggling greatly to find employment, as they likely have the most to gain by moving compared to their already diminished prospects at home. Indeed, the stark contrast between the OLS and IV estimates appears to support this interpretation.

While this study looks at one specific outcome, namely the time to re-employment, these short-term considerations are not the only ones that the unemployed face in their locational decision. Studies such as Kennan and Walker (2010) have addressed the importance of lifetime earnings expectations in the migration decision, although their study does not explicitly address the unemployed. Thus, further study is warranted to determine whether the unemployed in particular appear to be motivated by these longer-term income and career prospects. After all, different priorities may guide the migration decision for the unemployed than for the population at large, but the results in the present analysis suggest that shortening their unemployment spells appears to be a key consideration.

Appendix A: Regression Results for Demographic Subsamples

Table A.1: Results by Sex

	(1)	(2)	(3)	(4)
House Price measure	**Neg_equity**	**Neg_equity**	**Δln(HPI)**	**Δln(HPI)**
Subsample	**Female**	**Male**	**Female**	**Male**
	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)
HousePriceXOwn	**-0.001**	**-0.003*****	**0.005*****	**0.006*****
	(0.001)	**(0.001)**	**(0.001)**	**(0.001)**
Own	-0.000	0.009*	-0.002	-0.011*
	(0.003)	(0.004)	(0.003)	(0.004)
HousePrice	-0.001	0.001	-0.002*	-0.002**
	(0.000)	(0.001)	(0.001)	(0.001)
N	˜3,900,000	˜3,400,000	˜3,900,000	˜3,400,000

* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$

Notes: Sample consists of approximately 7.3 million ACS individuals i, either household head or spouse age 25-54, who are observed in one of 356 CBSAs j during a given quarter t in 2002q1-2012q4. All regressions include the following sets of controls: Ownership status X CBSA, CBSA X time period, Ownership status X time period. Individual controls for sex, age, race, education, marital status, presence of children, and residential tenure length are also included. Standard errors are clustered by CBSA.

Table A.2: Results by Race

House Price measure **Subsample**	**(1)** **Neg_equity** **Nonwhite**	**(2)** **Neg_equity** **White**	**(3)** **Δln(HPI)** **Nonwhite**	**(4)** **Δln(HPI)** **White**
	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)
HousePriceXOwn	**-0.002***	**-0.0021*****	**0.005*****	**0.006*****
	(0.001)	**(0.001)**	**(0.001)**	**(0.001)**
Own	-0.005	-0.004	-0.007	-0.006*
	(0.005)	(0.003)	(0.005)	(0.003)
HousePrice	0.000	0.001	-0.002	-0.003**
	(0.001)	(0.0001)	(0.001)	(0.001)
N	˜1,500,000	˜5,800,000	˜1,500,000	˜5,800,000

* p $\leq$ 0.05, ** p $\leq$ 0.01, *** p $\leq$ 0.001

Notes: Sample consists of approximately 7.3 million ACS individuals i, either household head or spouse age 25-54, who are observed in one of 356 CBSAs j during a given quarter t in 2002q1-2012q4. All regressions include the following sets of controls: Ownership status X CBSA, CBSA X time period, Ownership status X time period. Individual controls for sex, age, race, education, marital status, presence of children, and residential tenure length are also included. Standard errors are clustered by CBSA.

Table A.3: Results by Age Group

House Price measure **Subsample**	**(1)** **Neg_equity** **25-34**	**(2)** **Neg_equity** **35-44**	**(3)** **Neg_equity** **45-54**	**(4)** **Δln(HPI)** **25-34**	**(5)** **Δln(HPI)** **35-44**	**(6)** **Δln(HPI)** **45-54**
	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)
HousePriceXOwn	**-0.002***	**-0.002***	**-0.001**	**0.008*****	**0.003****	**0.002**
	(0.001)	**(0.001)**	**(0.001)**	**(0.001)**	**(0.001)**	**(0.002)**
Own	-0.010*	-0.002	0.001	-0.012**	-0.003	0.006
	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)	(0.000)
HousePrice	0.001	0.001	0.000	-0.003*	-0.001	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
N	˜2,000,000	˜2,500,000	˜2,800,000	˜2,000,000	˜2,500,000	˜2,800,000

* p $\leq$ 0.05, ** p $\leq$ 0.01, *** p $\leq$ 0.001

Notes: Sample consists of approximately 7.3 million ACS individuals i, either household head or spouse age 25-54, who are observed in one of 356 CBSAs j during a given quarter t in 2002q1-2012q4. All regressions include the following sets of controls: Ownership status X CBSA, CBSA X time period, Ownership status X time period. Individual controls for sex, age, race, education, marital status, presence of children, and residential tenure length are also included. Standard errors are clustered by CBSA.

Table A.4: Results by Earnings Quartile

House Price measure Subsample	(1) Neg_equity 1st Quartile	(2) Neg_equity 2nd Quartile	(3) Neg_equity 3rd Quartile	(4) Neg_equity 4th Quartile	(5) Δln(HPI) 1st Quartile	(6) Δln(HPI) 2nd Quartile	(7) Δln(HPI) 3rd Quartile	(8) Δln(HPI) 4th Quartile
	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)
HousePriceXOwn	**0.001**	**-0.001**	**-0.002***	**-0.002****	**0.005***	**0.004***	**0.005****	**0.004*****
	(0.001)	**(0.001)**	**(0.001)**	**(0.001)**	**(0.002)**	**(0.002)**	**(0.001)**	**(0.001)**
Own	-0.003	-0.002	-0.005	0.010**	-0.005	-0.003	-0.006	-0.012**
	(0.007)	(0.004)	(0.004)	(0.004)	(0.007)	(0.004)	(0.004)	(0.004)
HousePrice	0.001	0.000	0.000	0.001	-0.003**	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
N	~1,800,000	~1,800,000	~1,800,000	~1,800,000	~1,800,000	~1,800,000	~1,800,000	~1,800,000

* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$

Notes: Sample consists of approximately 7.3 million ACS individuals i, either household head or spouse age 25-54, who are observed in one of 356 CBSAs j during a given quarter t in 2002q1-2012q4. All regressions include the following sets of controls: Ownership status X CBSA, CBSA X time period, Ownership status X time period. Individual controls for sex, age, race, education, marital status, presence of children, and residential tenure length are also included. Standard errors are clustered by CBSA.

Table A.5: Results by Marital Status

	(1)	(2)	(3)	(4)
House Price measure	**Neg_equity**	**Neg_equity**	**Δln(HPI)**	**Δln(HPI)**
Subsample	**Married**	**Unmarried**	**Married**	**Unmarried**
	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)
HousePriceXOwn	**-0.003****	**-0.002****	**0.005*****	**0.006*****
	(0.001)	**(0.001)**	**(0.001)**	**(0.001)**
Own	-0.007	-0.001	-0.009*	-0.003
	(0.004)	(0.003)	(0.004)	(0.003)
$HousePrice$	0.001	0.001	-0.003*	-0.002*
	(0.001)	(0.001)	(0.001)	(0.001)
N	˜5,300,000	˜2,000,000	˜5,300,000	˜2,000,000

* p $\leq$ 0.05, ** p $\leq$ 0.01, *** p $\leq$ 0.001

Notes: Sample consists of approximately 7.3 million ACS individuals i, either household head or spouse age 25-54, who are observed in one of 356 CBSAs j during a given quarter t in 2002q1-2012q4. All regressions include the following sets of controls: Ownership status X CBSA, CBSA X time period, Ownership status X time period. Individual controls for sex, age, race, education, marital status, presence of children, and residential tenure length are also included. Standard errors are clustered by CBSA.

Table A.6: Results by Presence of Children

	(1)	(2)	(3)	(4)
House Price measure	**Neg_equity**	**Neg_equity**	**Δln(HPI)**	**Δln(HPI)**
Subsample	**Children**	**No Children**	**Children**	**No Children**
	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)
HousePriceXOwn	**-0.001***	**-0.003*****	**0.003****	**0.007*****
	(0.000)	**(0.001)**	**(0.001)**	**(0.001)**
Own	-0.004	-0.005	-0.005**	-0.008***
	(0.004)	(0.003)	(0.004)	(0.003)
$HousePrice$	0.000	0.001	-0.001	-0.003*
	(0.001)	(0.001)	(0.001)	(0.001)
N	˜4,300,000	˜3,300,000	˜4,300,000	˜3,300,000

* p $\leq$ 0.05, ** p $\leq$ 0.01, *** p $\leq$ 0.001

Notes: Sample consists of approximately 7.3 million ACS individuals i, either household head or spouse age 25-54, who are observed in one of 356 CBSAs j during a given quarter t in 2002q1-2012q4. All regressions include the following sets of controls: Ownership status X CBSA, CBSA X time period, Ownership status X time period. Individual controls for sex, age, race, education, marital status, presence of children, and residential tenure length are also included. Standard errors are clustered by CBSA.

Table A.7: Results by Education Group

House Price measure Subsample	(1) **Neg_equity** **Less than HS**	(2) **Neg_equity** **HS**	(3) **Neg_equity** **Some Coll**	(4) **Neg_equity** **Coll +**	(5) **Δln(HPI)** **Less than HS**	(6) **Δln(HPI)** **HS**	(7) **Δln(HPI)** **Some Coll**	(8) **Δln(HPI)** **Coll +**
	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)	β/(s.e.)
HousePriceXOwn	0.001	-0.001	-0.002*	-0.002*	-0.002	0.003	0.006***	0.006***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.003)	(0.002)	(0.002)	(0.001)
Own	-0.000	-0.013**	-0.001	-0.002	0.000	-0.014*	-0.004	-0.005
	(0.000)	(0.004)	(0.004)	(0.004)	(0.007)	(0.004)	(0.005)	(0.003)
HousePrice	-0.004*	0.000	0.000	0.001	0.005**	-0.000	-0.003***	-0.004***
	(0.001)	(0.001)	(0.001)	(0.0001)	(0.002)	(0.000)	(0.001)	(0.001)
N	~500,000	~1,600,000	~2,300,000	~2,900,000	~500,000	~1,600,000	~2,300,000	~2,900,000

* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$

Notes: Sample consists of approximately 7.3 million ACS individuals i, either household head or spouse age 25-54, who are observed in one of 356 CBSAs j during a given quarter t in 2002q1-2012q4. All regressions include the following sets of controls: Ownership status X CBSA, CBSA X time period, Ownership status X time period. Individual controls for sex, age, race, education, marital status, presence of children, and residential tenure length are also included. Standard errors are clustered by CBSA.

Bibliography

Aaronson, D. and J. Davis (2011): "How Much has House Lock Affected Labor Mobility and the Unemployment Rate," *Chicago Fed Letters*, 290.

Abowd, J., Stephens, B., Vilhuber, L., Andersson, F., McKinney, K., Roemer, M., and S. Woodcock (2005): "The LEHD Infrastructure and the Creation of the Quarterly Workforce Indicators," U.S. Census Bureau's Center for Economic Studies Technical Paper No. TP-2006-01.

Abowd, J. and L. Vilhuber (2011): "National Estimates of Gross Employment and Job Flows from the Quarterly Workforce Indicators with Demographic and Industry Detail," *Journal of Econometrics*, 161(1), 82-99.

Abowd, J. and L. Vilhuber (2012): "Did the Housing Price Bubble Clobber Local Labor Market Job and Worker Flows When it Burst," *American Economic Review*, 102(3), 589-593.

Abraham, K., Haltiwanger, J., Sandusky, K., and J. Spletzer (2009): "Exploring Differences in Employment between Household and Establishment Data," *Journal of Labor Economics*, 31(2), 129-172.

Angrist, J. (2001): "Estimation of Limited Dependent Variable Models with Dummy Endogenous Regressors: Simple Strategies for Empirical Practice," *Journal of Business and Economics*, 19(1), 2-16.

Angrist, J. and G. Imbens (1995): "Two-Stage Least Squares of Average Causal Treatment Effects in Models with Variable Treatment Intensity," *Journal of of the American Statistical Association*, 90, 431-42.

Bailey, A.J. (1994): "Migration and Unemployment among Young Adults," *Papers in Regional Science*, 73(3), 289-307.

Basker, E. (2003): "Education, Job Search and Migration," *University of Missouri Working Paper.*

Bijwaard, G. and G. Ridder (2005): "Correcting for Selective Compliance in a Re-employment Bonus Experiment," *Journal of Econometrics*, 125, 77-111.

Bitter, C. (2008): "Geography, Housing Prices, and Interregional Migration," *Dissertation submitted to the University of Arizona, Department of Geography.*

Borjas, G.(1991): "Immigration and Selection," in *Immigration, Trade and the Labor Market* ed. J. Abowd and R. Freeman, University of Chicago Press, 29-76.

Bureau of Labor Statistics. *Local Area Unemployment Statistics, 2009.* [Data file]. Available from www.bls.gov/lau/.

Cameron, A.C. and P.K. Trivedi (2005): "Microeconometrics: Methods and Applications," *Cambridge University Press.*

Cameron, G. and J. Muellbauer (1998): "The Housing Market and Regional Commuting and Migration Choices," *Scottish Journal of Political Economy*, 45(4), 420-46.

Chan, S. (2001): "Spatial Lock-in: Do Falling House Prices Constrain Residential Mobility,". *Journal of Urban Economics*, 49(3), 567-586.

Cheshire, P. and S. Sheppard (1995): "On the Price of Land and the Value of Amenities," *Economica*, 62(246), 247-267.

Cooper, D. (2013): "The Effect of Unemployment Duration on Future Earnings and Other Outcomes," *Federal Reserve Bank of Boston Working Paper Series*, 13-8.

CoreLogic, Inc. *Negative Equity Report 2010:Q4*, March 8, 2011.

DaVanzo, J. (1978): "Does Unemployment Affect Migration? Evidence from Micro Data," *Review of Economics and Statistics*, 60(4), 504-514.

DaVanzo, J. (1983): "Return Migration in the United States: Who moves back and who moves on?" *The Review of Economics and Statistics*, 65(4), 552-59.

Davies, P., Greenwood, M. and H. Li (2001): "A Conditional Logit Approach to U.S. State-to-State Migration," *Journal of Regional Science*. 41(2), 337-360.

Davis, S. and J. Haltiwanger (2014): "Labor Market Fluidity and Economic Performance," *NBER Working Papers* 20479, National Bureau of Economic Research, Inc.

Deaton, A. (1997): "The Analysis of Household Surveys: A Microeconometric Approach to Development Policy," *Johns Hopkins University Press* London and Baltimore.

Decker, R., Haltiwanger, J., Jarmin, R. and J. Miranda (2014): "The Role of Entrepreneurship in US Job Creation and Economic Dynamism," *Journal of Economic Perspectives*,28(3), 3-24.

Demyanyk, Y. and O. Van Hemert (2011): "Understanding the Subprime Mortgage Crisis," *Review of Financial Studies*, 24(6), 1848-1880.

Dennett, J. and A. Modestino (2011): "Are American Homeowners Locked in their Houses: The Impact of Housing Market Conditions on State-to-State Migration," *Federal Reserve Bank of Boston Working Paper* 12-1.

Engelhardt, G. (2003): "Nominal Loss Aversion, Housing Equity Constraints, and Household Mobility," *Journal of Urban Economics*, 53(1), 171-195.

Farber, H. (2010): "Job Loss in the Great Recession: Historical Perspectives from the Displaced Workers Survey 1984-2010," *NBER Working Paper* 17040.

Ferreira, F., Gyourko, J. and J. Tracy (2010): "Housing Busts and Household Mobility," *Journal of Urban Economics*, 68, 34-45.

Ferreira, F., Gyourko, J. and J. Tracy (2011): "Housing Busts and Household Mobility: An Update," *Federal Reserve Bank of New York Working Paper* Staff Report 526.

Gabriel, S., Shack-Marquez, J. and W. Wascher (1992): "Regional house-price dispersion and interregional migration," *Journal of Housing Economics*, 2(3), 235-25.

Genesove, D. and C. Mayer (1997): "Equity and Time to Sale in the Real Estate Market," *American Economic Review*, 87(3), 255-269.

Genesove, D. and C. Mayer (2001): "Loss-aversion and Seller Behavior: Evidence from the Housing Market," *Quarterly Journal of Economics*, 116(4), 1233-60.

Gerardi, K., A.H. Shapiro and P. Willen (2008): "Subprime Outcomes: Risky Mortgages, Homeownership Experiences, and Foreclosures," *Federal Reserve Bank of Boston Working Paper* 07-15.

Ghent, A. and M. Kudlyak (2009): "Recourse and Residential Mortgage Default: Theory and Evidence from the U.S.," *Federal Reserve Bank of Richmond Working Paper* 09-10 .

Goss, E., Paul, C., and A. Wilhite (1994): "Duration of Unemployment: Geographic Mobility and Selectivity Bias," *Review of Regional Studies*, 24, 127-142.

Graves, P. (1983): "Migration with a composite amenity': the role of rents," *Journal of Regional Science*, 23(4), 541-546.

Graves, P. and P. Linneman (1979): "Household Migration: Theoretical and Empirical Results," *Journal of Urban Economics*, 6, 383-404.

Graves, P. and P. Mueser (1995): "Examining the role of economic opportunity and amenities in explaining population redistribution," *Journal of Urban Economics*, 37(2), 176-200.

Greenwood, M. (1975): "Research on Internal Migration in the United States: A Survey," *Journal of Economic Literature*, 13, 397-433.

Greenwood, M. (1997): "Internal Migration in Developed Countries," in "Handbook of Population and Family Economics," ed. M.R. Rosenzweid and O. Stark, 647-720. *Elsevier Science*, New York.

Gregory, M. and R. Jukes (2001): "Unemployment and Subsequent Earnings: Estimating Scarring among British Men 1984-94," *Economic Journal*, 111(475), 607-625.

Harris, J. and M. Todaro. (1970): "Unemployment and Development: A Two-Sector Analysis," *American Economic Review*, 60(1), 126-42.

Haurin, D. and R.J. Haurin (1988): "Net Migration, Unemployment, and the Business Cycle," *Journal of Regional Science*, 28(2), 239-254.

Heckman, J. (1979): "Sample Selection Bias as Specification Error."*Econometrica*, 47, 153-161.

Herzog, H. and A. Schlottmann (1981): "Employment Status and the Decision to Migrate," *Review of Economics and Statistics*, 63, 590-598.

Herzog, H.W., Schlottman, A. and T.P. Boehm (1993): "Migration as Spatial Job-Search: A Survey of Empirical Findings," *Regional Studies*, 27(4), 327-340.

Hoynes, H., Miller, D., and J. Schaller (2012): "Who suffers during recessions?" *Journal of Economic Perspectives*, 26(3), pp. 27-48.

Hunt, G., and R. Mueller (2004): "North American Migration: Returns to Skill, Border Effects,and Mobility Costs," *The Review of Economics and Statistics*, 86, pp. 988-1007.

Hyatt, H. and E. McEntarfer (2012): "Job-to-Job Flows in the Great Recession," *American Economic Review*, 102(3), 580-83.

Hyatt, H.. and J. Spletzer (2013): "The Recent Decline in Employment Dynamics," *IZA Journal of Labor Economics*, 2(5).

Jackman, R. and S. Savouri (1992): " Regional migration versus regional commuting : the identification of housing and employment flows," *Scottish Journal of Political Economy*, 39(3), 272

Kahnneman, D. and A. Tversky (1979): "Prospect Theory: An Analysis of Decision Under Risk,". *Econometrica*, 47, 263-291.

Katz, L. (2010): "Hearing on Long-Term Unemployment: Causes, Consequences, and Solutions," Testimony before the Joint Economic Committee of U.S. Congress. April 29.

Katz, L. (2010, Jan. 10): "A Nation of Hunkered Down Homebodies", *The New York Times*, editorial.

Kennan, J. and J. R. Walker (2011): "The Effect of Income on Individual Migration Decisions," *Econometrica*, 79(1), 211-51.

Knapp, T. and P. Graves (1989): "On the Role of Amenities in Models of Migration and Regional Development," *Journal of Regional Science*, 29(1), 71-87.

Knight, J. and S. Li (2006): "Unemployment duration and earnings of re-employed workers in urban China," *China Economic Review*, 17(2), 103-119.

Lewis, A. (1954): "Economic Development with Unlimited Supplies of Labor," *The Manchester School*. 22, 139-191.

Long, L. (1991): "Residential Mobility Differences Among Developed Countries," *International Region Sciences Review*, 14(2), 133-147.

Maddala, G. (1983): "Limited Dependent and Qualitative Variables in Econometrics," *Cambridge University Press*, Cambridge.

Malamud, O. and A. Wozniak (2012): "The Impact of College Education on Geographic Mobility," *Journal of Human Resources*, 47(4), 913-950.

McCall, B. (1996): "Unemployment Insurance Rules, Joblessness, and Part-Time Work,". *Econometrica*, 64(1), 647-682.

Meyer, B. (1986): "Semiparametric Estimation of Hazard Models," *MIT Working Paper*.

Meyer, B. (1990): "Unemployment Insurance and Unemployment Spells," *Econometrica*, 58(4), 757-782.

Mincer, J. (1978): "Family Migration Decisions," *Journal of Political Economy*, 86, 749-773.

Molho, I. (1986): "Theories of Migration: A Review," *Scottish Journal of Political Economy*, 33, 396-419.

Molloy, R., Smith, C. and A. Wozniak (2011): "Internal Migration in the United States," *Journal of Economic Perspectives*, 25(3), 173-196.

Molloy, R, Smith, C..and A. Wozniak (2014): "Declining Migration within the U.S.: The Role of the Labor Market," *IZA Discussion Paper* 8149, Institute for the Study of Labor (IZA).

Molloy, R.and A. Wozniak. (2011): "Labor Reallocation over the Business Cycle: New Evidence from Internal Migration," *Journal of Labor Economics*, 29(4), 697-739.

Nakosteen, R., Westerlund, O., and M. Zimmer (2008): "Migration and Self-Selection: Measuring Earnings and Latent Characteristics," *Journal of Regional Science*, 48(4), 769-788.

Office of Management and Budget. *OMB Bulletin No. 10-02: Update of Statistical Area Definitions and Guidance on Their Uses,* Dec. 1, 2009.

Pekkala S. and H. Tervo (2002): "Unemployment and Migration: Does Moving Help?" *Scandinavian Journal of Economics*, 104(4), 621-639.

Quigley, J.M. (1987): "Interest Rate Variations, Mortgage Prepayments and Household Mobility," *Review of Economics and Statistics*, 69, 636-643.

RealtyTrac. *U.S. Foreclosure Market Report: Year-End 2010*, Jan. 13, 2011.

Rogerson, P. and R. McKinnon (2005): "A Geographic Model of Job Search, Migration and Unemployment," *Papers in Regional Science*, 48(1), 89-102.

Sandell, S. (1977): "Women and the Economics of Family Migration," *Review of Economics and Statistics*, 59, 406-411.

Sato, Y. (2004): "Migration, Frictional Unemployment, and Welfare Improving Policies," *Journal of Regional Science*, 44, 773-793.

Schaeffer, P. (1988): "Family migration in an uncertain environment," *Socio-Economic Planning Sciences*, 22(5), 221-227.

Schmitt, J. and C. Warner (2011): "Deconstructing Structural Unemployment," *Center for Economic and Policy Research Papers* .

Schulhofer-Wohl, S. (2012): "Negative Equity does not reduce homeowners' mobility," *Federal Reserve Bank of Minneapolis Quarterly Review* 35(1), 2-14..

Shimer, R. (2012): "Reassessing the Ins and Outs of Unemployment," *Review of Economic Dynamics*. 15(2), 127-148.

Shumway, J.M. (1993): "Factors influencing Unemployment Duration with a special emphasis on Migration," *Papers in Regional Science*, 72(2), 159-176.

Sjaastad, L. (1962): "The Costs and Returns of Human Migration," *Journal of Political Economy* ,70(5), 80-93.

Staiger, D. and J. Stock (1997): "Instrumental Variables regression with weak instruments," *Econometrica*, 65(3), 557-586.

Stein, J. (1995): "Prices and Trading Volume in the Housing Market: A Model with Down-Payment Effects," *Quarterly Journal of Economics*, 110, 379-406.

Thiebout, C. (1956): "A Pure Theory of Local Expenditures," *Journal of Political Economy*, 64(5), 416-424.

Todaro, M. (1969): "A Model of Labor Migration and Urban Unemployment in Less Developed Countries," *The American Economic Review*, 59(1), 138-148.

Valetta, R. (2013): "House Lock and Structural Unemployment," *Labour Eocnomics* , 25(3):86-97.

Vijverberg, W. (1995): "Dual Selection Criteria with Multiple Alternatives: Migration Work Status and Wages," *International Economic Review*, 36, 159-185.

Wozniak, A. (2010): "Are College Graduates More Responsive to Distant Labor Market Opportunities," *Journal of Human Resources*, 45(4), 944-970.

Zenou, Y. (2011): "Rural-urban migration and unemployment. Theory and policy implications," *Journal of Regional Science*, 51, 65-82.

Zillow Group. *Zillow Home Value Index, Dec 2012*. [Data file]. Available from Zillow.com/research/data.