NBER WORKING PAPER SERIES

RESIDENTIAL MOBILITY AND MORTGAGES

Sewin Chan

Working Paper No. 5181

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 1995

This paper was prepared for the NBER Conference on Public Policy and the Housing Market. I thank Andrew Caplin, Michael Cragg, Charles Freeman and Joseph Tracy for their guidance and encouragement, and Rosanne Altshuler, Omar Azfar, Ronald Miller, Ali Moazami, Jörg Oechssler and Ken Xue for their help and suggestions. I also thank Jan Brueckner, Henry Buist, Patric Hendershott and other participants of the NBER Conference for their comments and discussion. Chemical Bank's advice, assistance and tremendous knowledge of mortgage markets have been indispensable to this project and their enthusiasm for this kind of research is most encouraging. The generosity of Harold Free and Phil O'Hara in providing access to Public Data Corporation's automated public records database is gratefully acknowledged. The expert assistance of Sue Clay is especially appreciated. This paper was prepared as part of the NBER Study on Housing Dynamics and was presented at the NBER Study Conference. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

NBER Working Paper #5181
July 1995

RESIDENTIAL MOBILITY AND MORTGAGES

ABSTRACT

Mortgage applications are a detailed and accurate source of household information that is verified by underwriters, making it a more accurate data source than self-reported survey answers. This paper discusses how mortgage data can be applied to areas of economics outside mortgage finance. As a supplement to variables from the application form, the self-selection of mortgage points is used to infer expected mobility. A duration model of housing spells is estimated, and the points indicator is shown to be highly significant in predicting mobility for low loan-to-value borrowers. The findings demonstrate the potential fruitfulness of using this new data source.

Sewin Chan
Department of Economics
Columbia University
New York, NY 10027

1. Introduction

The residential housing market is a very important part of the economy that impinges on many aspects of consumer behavior. In the U.S., some 65% of households own their homes (US Bureau of the Census), and for the vast majority, their dwelling is the single largest component of their wealth. The substantial cost of residential properties compared with household income, and the tax deductibility of mortgage interest, means that most homeowners use mortgages to finance their purchase.

Tremendous information is collected as part of the mortgage applications process. The data constitutes one of the most accurate sources available to researchers: it is verified by underwriters rather than self-reported by survey. Moreover, the self-selection nature of the mortgage contract menu forces households to reveal some of their preferences and expectations; in particular, regarding mobility and interest rates.

The potential value of mortgage applications data is far-reaching. It can be used to study important issues in economics, such as mobility and location choice. Somewhat surprisingly, the accuracy of mortgage data has rarely been exploited in examining these issues before, and this paper will elaborate on how we can further expand our knowledge in these areas, which has so far been garnered mostly from survey data.

A sample of mortgage loans from Chemical Bank is presented. It expands on the dataset used by Caplin, Freeman and Tracy (1993) by including applications data: detailed demographic, location, asset and income information. The data has also been

merged with loan performance information that tells us whether a borrower moved or refinanced the mortgage.

As well as containing all the variables that can be read directly off the application form, the self-selection of borrowers into different mortgage contracts allows us to infer their expected mobility. When taking a mortgage, borrowers can choose to pay "points", an initial lump-sum fee, in exchange for a reduction in the interest rate on their loan. Those who plan to move in the near future will not have an incentive to pay these points because most mortgages are non-assumable. On the other hand, borrowers who plan to hold their mortgage for many years will have an incentive to trade points for a reduction in their rate. Points are therefore an indicator of expected mobility.

The concept is illustrated with a housing spell duration model estimated from the applications data, first using standard demographic, income and asset variables, and then utilizing the points indicator. One would expect that absent any unanticipated constraints to mobility, the points indicator would be very significant in explaining mobility, and this is indeed confirmed for low loan-to-value (LTV) borrowers. For high LTV borrowers, the recent unanticipated decline in property values may have led to insufficient funds for a downpayment on a new home, and so points are not a good predictor of mobility for this group.

The analysis is illustrative of the potential gains from using mortgage applications data to study areas of economics outside of mortgage finance. Its accuracy and wide applicability clearly justify the efforts required in collecting data of this kind.

The remainder of the paper is structured as follows: the next section discusses the potential contribution of mortgage applications data. Section 3 shows how mobility expectations can be inferred from the number of points paid, both for fixed and adjustable rate mortgages. Section 4 presents the dataset to be used in the analysis, and section 5 presents the results. Finally, section 6 concludes.

2. The Contribution of Mortgage Applications Data

Mortgage applications data are a promising source of information for economic research because housing is a component concern in many economic decisions, including employment, schooling, the choice of local amenities and local taxation. A major contribution of mortgage data is the accuracy of the variables compared with survey data. Retrospective questions, sometimes asking respondents to refer to an event that occurred a year ago, are subject to the distorting effects of memory loss and rationalization. With mortgage data, these variables come directly from the application system that is used in underwriting and that generates the necessary documentation for the loan to be sold on secondary markets. The data are verified by experienced underwriters and thus reliability is of the highest order possible.

The standard Fannie Mae application form elicits much information that is directly relevant to the analysis of consumer behavior. It asks the borrower for seven years of location history, and at least two years of this information is verified by underwriters,

either by contacting the landlord (for renters) or examining the deed (for owners). Two years of employment history is required, and the employer is contacted for verification. Demographic variables include the applicant's age, years of school, marital status and number of children. Income, assets and liabilities are verified by paystubs, tax returns, bank statements, and credit bureau reports. The exact address of the property to be mortgaged and its appraised value are also available. This is in addition to variables that specify the nature of the mortgage contract, such as the interest rate, the adjustability of the interest rate, the loan balance and the number of points paid.

Because housing is a spatial good, issues related to mobility and location choice are natural candidates for investigation with this data. The benefits and contributions of mortgage data will be illustrated and contrasted with previous efforts in this field.

2.1 Mobility and Location Choice

Mobility is important at a macroeconomy level because it plays a predominant role in regional adjustment. Blanchard & Katz (1992), among others, demonstrated that the principle mechanism of regional economic adjustment in the U.S. is labor mobility rather than job creation or job migration. On a consumer welfare level, households typically have to move to achieve their desired level and location of housing consumption, and this desire may shift over time because of changing jobs and household composition. Similarly for the allocation of local public goods: movement over the lifecycle induces changes in preferences for locally provided services such as education and healthcare. Efficiency of local public expenditures in the Tiebout (1956) sense can only be achieved if mobility is

unhindered. Thus, for all these reasons, a better understanding of residential mobility is a worthwhile goal.

Previous studies of mobility have relied on several sources of survey data. On an aggregate level, the US Census includes a question asking individuals about their place of residence five years earlier. Mobility is then based on comparing previous and current addresses. The Current Population Survey also asks retrospective residence questions, of both one and five years. Many studies have been done using these data, taking advantage of their large sample size and geographic coverage. However, there are problems with missing multiple moves and return migration that have been well documented (see Lichter and De Jong (1990)).

Several panel surveys, including the Panel Survey of Income Dynamics (PSID) and the National Longitudinal Surveys, have the advantage of being able to track the same households through time as they move. Their detailed questions regarding work and income make them very useful in studying the determinants of mobility especially in relation to the labor market (for example, Bartel (1979) and Borjas, Bronars & Trejo (1992)). However, they suffer from both small sample size and sample attrition that can lead to bias, as non-respondents are disproportionately mobile. The Annual Housing Survey (AHS) is a panel that tracks housing units rather than households, and therefore cannot be used directly as a panel in analyzing mobility. Gronberg and Reed (1992) propose an algorithm to circumvent this problem.

Each of these data sources has its advantages and disadvantages, and each has been used extensively to analyze a wide variety of questions. Regarding mobility, the

dramatic fluctuations in house prices and interest rates over the past decade has prompted much research concerning their effects on the mobility of homeowners.

Quigley (1987) uses PSID data to investigate the effects of mortgages at more favourable than current interest rates. He finds that these effects are quite large; having a fixed rate mortgage at below market rates significantly increases duration time in the home. Harmon and Potepan (1988) come to a similar conclusion when they examine how changes in adjustment costs, as well as housing demand factors such as income, influence mobility. Potepan (1987) argues that higher market interest rates make home improvements more attractive relative to moving. Using the PSID, he shows that the probability of making home improvements increases with the current interest rate and decreases with income. Henderson and Ioannides (1989) use the PSID and find that wealthy and more educated families tend to be more mobile. Gronberg and Reed (1992) find similar results using the AHS.

Nakagami and Pereira (1991) develop a theoretical model that focuses on the optimal timing of "trading up" and how this is affected by macroeconomic variables. They predict that appreciation leads to faster tradeup and hence increased mobility, and that rising interest rates reduces the incentive to tradeup and therefore lowers mobility. Kiel (1993) finds support for this theory using the AHS: middle aged homeowners with more than 5 years tenure were more likely to move if their home experienced high appreciation.

The recent decline in regional house prices and the general fall in interest rates has provided researchers with a new environment in which to analyze mobility issues and a new set of questions to answer, since it is not clear that fluctuations in house prices and

interest rates will have symmetric effects in both directions. Mortgage data can and should play a role in the analysis of these issues.

Stein (1993) has constructed a theoretical model that illustrates how the effects of downpayment constraints on repeat buyers can reduce mobility. When prices fall, households who must rely on the sale of their current home for the downpayment on a new home may be prevented from moving. Constrained households may "fish" for a buyer: list the house for sale at a price above market in the hope that a buyer will be found. Genesove and Mayer (1993) use a sample of condominiums listed for sale in Boston to test this hypothesis. Their results show that high loan-to-value (LTV) properties are much less likely to sell, and that if a sale occurs, the transaction price is significantly higher than for properties with low LTVs.

With mortgage data, we can examine this question at the household level. Are high LTV and low asset households constrained from moving? Are people in properties with larger price falls more constrained from moving? The self-selection nature of mortgage contracts allows us to determine a household's mobility expectation (as detailed in section 3), and we can see which households expected to move and were prevented from doing so by unanticipated shocks such as house price movements.

Mortgage applications data also have the unique characteristic of being able to tell us exactly where households want to move. All applications whether rejected or accepted by the bank can be used, and a several hundred dollar application fee ensures that this preference is serious. Both the household's address at the time of the application and the desired address are available, and this information can be matched with data on the

characteristics of those particular areas. Tiebout type models can be tested, making use of the detailed income, asset and demographic information.

2.2 Difficulties in Obtaining Mortgage Applications Data

Mortgage data are seldom used for empirical analysis outside of mortgage finance and even within this field, all the information available is rarely fully exploited. Mortgage data currently available for research are from mortgage pools, surveys conducted by the National Association of Realtors, panel surveys such as the AHS and special waves of the PSID, directly from financial institutions that originate mortgages, and from the Government Sponsored Agencies.

Pool data suffer from the problems of aggregation. In general, data on the individual loans are not available; only average behavior and characteristics can be observed. Surveys suffer from the effects of memory loss and rationalization mentioned earlier. Mortgage applications data obtained directly from the financial institutions that originated them are much more reliable. A strict subset of this information is passed onto the servicer and the Agencies, and so the most complete set lies with the originator of the mortgage.

Applications are filled in by all potential borrowers, and today, most originators facilitate processing by entering the entire application into electronic form. However, the task of data collection is non-trivial. First and foremost, institutional cooperation is essential. Besides obtaining access to commercially valuable information, ties and

relationships need to be developed with the originator so that the complexities and idiosyncrasies of each origination system can be explained and understood.

Secondly, information is generated as a byproduct of the mortgage origination business. The systems used are excellent for the purposes of underwriting and servicing, but are not conducive to the needs of researchers. Many months were spent painstakingly constructing a useable dataset from backup tapes of the origination system and linking it to information provided by the servicer of these mortgages.

It is a shame that mortgage data are not automatically collected in a more useable way. The fact that most mortgages are conforming means that there is a huge potential for gathering data from different institutions and pooling it into one giant panel database. We would then be able to track a household's behavior even if it changed financial institutions.

3. Inferring Expected Mobility from Points

As a supplement to the variables that can be read directly from the application form, a household's expected mobility can be inferred from the choice of mortgage contract. Mortgage banks typically allow customers to reduce the interest rate on a loan by paying discount points, an upfront cost that is expressed as a percentage of the loan amount. Clearly, the value of this discount depends on how long the mortgage is held. This rate-point trade-off and its relation to expected tenure has been well documented in

the housing finance literature; for example, Stone & Zissu (1990) and Yang (1992) give theoretical discussions of the tradeoff, and Brueckner (1994) provides some evidence confirming this behavioral hypothesis.

As an illustration, suppose interest rates are 9% for a 30 year fixed-rate mortgage (FRM) with zero points and 8.5% with two points. With a loan balance of $100,000, a payment of $2000 (two points) will reduce the interest rate by 0.5 percentage points. For the moment, assume that the borrower does not have expectations of falling interest rates, i.e., there is no expectation of refinancing the mortgage. If the borrower expects hold the mortgage for 10 years, then she should pay points; if she plans to move within two years, then she should not pay points. The break-even holding period where she would be indifferent between paying points or not is approximately 60 months, depending on the discount rate and the marginal tax rate (both points and interest payments are tax deductible). A lower discount rate or marginal tax rate would reduce the break-even period.

Complications arise in a falling rate expectation environment that may induce non-movers to wait for rates to fall and then refinance, rather than reduce the rate with points. Suppose the borrower contemplating the FRM at 9% with no points expects to stay for 10 years, then she should reduce the rate by paying points. However, she also expects rates to fall by 1 percentage point every year for the next 2 years until it reaches 7%. In this case, it would be better to pay no points and refinance after two years. Table 3.1 summarizes the point choice of borrowers with different rate and mobility expectations.

Table 3.1 Choice of points for movers and non-movers				
Interest Rate	Fixed rate mortgages		Adjustable rate mortgages	
Expectations	Movers	Non-movers	Movers	Non-movers
Rising	no points	points	no points	points
Constant	no points	points	no points	no points
Falling	no points	no points	no points	no points

For adjustable rate mortgage (ARMs), the point-rate tradeoff is more complicated. ARMs come in many varieties, the most common being the 1-year and the 3-year, where the time period refers to the time between interest rate adjustments and the relevant index. The interest rate is set using an index (for example, a treasury bill rate) plus a margin (for example, 2.75 percentage points) with maximum upward adjustment caps (for example, 2 percentage points each adjustment period and 6 percentage points over the life of the loan). Lenders also offer borrowers a discount or "teaser" rate, so that the initial rate in the period before the first adjustment is 1 to 2 percentage points below the index-plus-margin rate[1].

The trade-off between points and rate for ARMs depends crucially on interest rate expectations. As a rule, each point paid on an ARM reduces this initial rate by 0.5 percentage points for 1-year ARMs and 0.25 percentage points for ARMs with longer adjustment periods. The benefit from paying points can extend beyond this initial period because of the interest rate caps.

[1]However, under Fannie Mae guidelines, a potential borrower's debt-income ratio is calculated on the basis of the current index-plus-margin rate, not the teaser.

In a rising rate environment, the adjustment period cap may be binding for several periods. For example, consider a 1-year ARM with 2% and 6% adjustment-period and lifetime caps, which costs 6.5% with no points and 5.5% with two points, as shown in Table 3.2. The index-plus-margin rate stands at 8%. If market rates rise by 1% each year for the next three years, the index-plus-margin rate will increase to 9%, 10% and 11% after years 1, 2 and 3 respectively; but the actual rate paid (with no points) will be capped at 8.5% for the first year and then move with the index-plus-margin at 10% and 11% thereafter.

Table 3.2. Example of a 1-year ARM with rising interest rate expectations

	index + margin	ARM rate with zero points	ARM rate with two
initial	8.0	6.5	5.5
after year 1	9.0	8.5 (hit periodic cap)	7.5 (hit periodic cap)
after year 2	10.0	10.0	9.5 (hit periodic cap)
after year 3	11.0	11.0	11.0
after year 4	12.0	12.0	11.5 (hit lifetime cap)
after year 5	13.0	12.5 (hit lifetime cap)	11.5

By paying two points, the borrower will be able to buydown his initial rate to 5.5% and thus the rates paid will be 7.5%, 9.5% and 11%. The break-even holding period where one is indifferent between zero and two points is approximately 2 years with this set of rate expectations and outcomes. If rates continue to rise after 3 years, the ARM will eventually hit the lifetime cap; for the zero point ARM, this is 12.5% and for the 2 point ARM this is 11.5%.

Suppose, on the other hand, that market rates remain constant over time so that the index-plus-margin is always 8%, as shown in Table 3.3. After one year, if zero points are paid, the rate will jump to 8%. If 2 points are paid, the rate will be capped at 7.5% and in the second year, the rate will rise to 8%.

Table 3.3. Example of a 1-year ARM with flat interest rate expectations

	index + margin rate	ARM rate with zero points	ARM rate with two
initial	8.0	6.5	5.5
after year 1	8.0	8.0	7.5 (hit periodic cap)
after year 2	8.0	8.0	8.0
after year 3	8.0	8.0	8.0
after year 4	8.0	8.0	8.0

It would never be profitable to pay points in this case; with falling interest rates, the case against paying points would be even stronger[2]. Thus, points paid are only a predictor of expected mobility in rising rate environments for ARMs, and a non-falling rate environment for FRMs, as shown in Table 3.1.

In a rising rate environment, the number of points paid can provide an indicator of how long the borrower expects to hold the mortgage. Figure 3.1, where T_1, T_2 and T_3 refer to the three break-even periods between paying 0 & 1 point, 1 & 2 points and 2 & 3 points, illustrates this. If the borrower expects to hold the mortgage less than T_1, she will pay zero points, between T_1 and T_2, she will pay one point, between T_2 and T_3, two points and beyond T_3 three points.

[2]ARMs are not subject to the same refinance complication as ARMs. Many ARM contracts provide an option for borrowers to convert their adjustable rates to the current fixed rate in exchange for a small fee, which is substantially lower than the costs of a refinance. There are constraints on the times at which this option can be exercised, but these restrictions are not severe.

Fig. 3.1. Present value cost of a mortgage against holding time

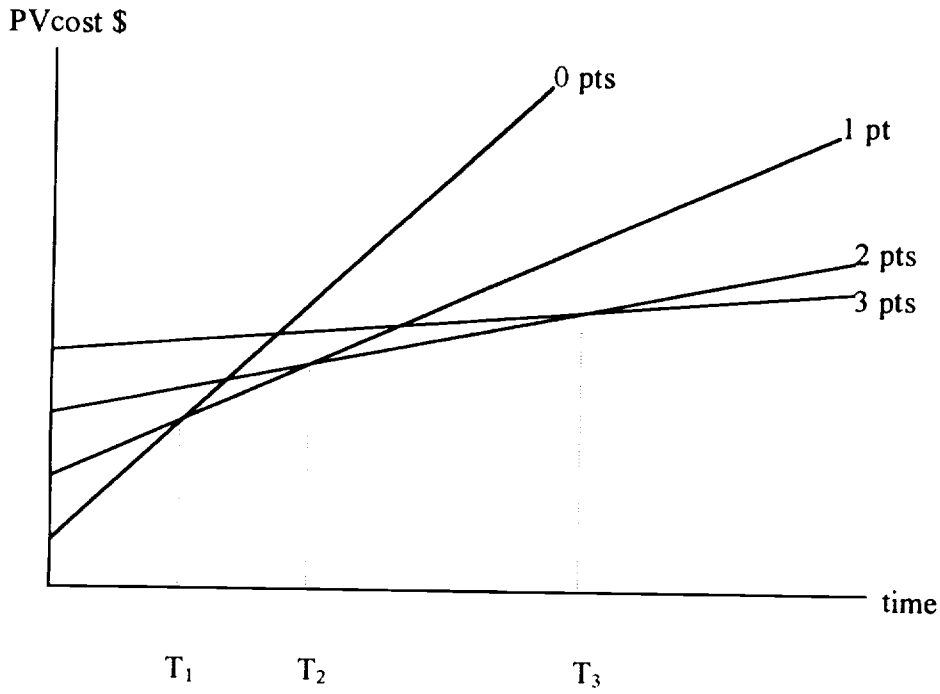
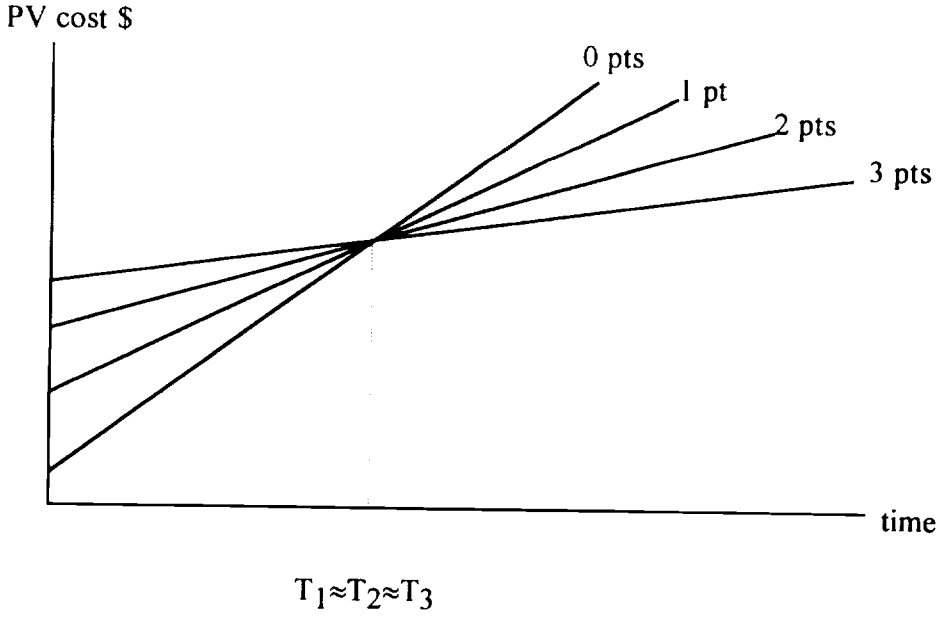


Fig. 3.2. Present value cost of a mortgage against holding time

for the current point-rate pricing policy



$T_1 \approx T_2 \approx T_3$

However, the pricing policy of a 0.25 percentage point lower rate for every point paid (adopted by many banks) leads to a situation where the incentive is to pay zero or the maximum number of points. For the 30 year FRM described above -- 9% for with zero points, 8.75% with 1 point, 8.5% with 2 points and 8.25% with 3 points -- the break-even holding periods T_1 , T_2 and T_3 are all equal to 5 years. Thus, the current pricing policy leads to a situation as depicted in Figure 3.2.

For holding periods shorter than, say, 5 years, the absolute difference between the present values will be small and so the borrower may choose to pay one point for risk aversion reasons. Therefore, the number of points paid may still be correlated with the length of the holding period.

In any case, points provide an indicator of the expected holding period. Those who expect to move in the near future should never pay points. Non-movers should pay points if they expect interest rates to rise. With constant or falling rates, ARM borrowers should not pay points whether moving or staying.

4. The Data

The data have been collected from a sample of 17,000 residential mortgages originated by Chemical Bank in four boroughs of New York City (The Bronx, Brooklyn, Manhattan and Queens) between 1989 and 1993. This is a very densely populated metropolitan area with a large dispersion of income and assets and a variety of property

types, including co-operatives (co-ops) and condominiums. All the data found on a mortgage application form, together with the fields used by underwriters are available. The information on the property and the borrower is both detailed and accurate because of verification by underwriters.

Chemical Bank is an A-paper lender, indicating that the borrowers in the sample have relatively good credit[3]. It also means that the "price" of a mortgage is not individually negotiated; rather, the borrower chooses from a menu of contracts that vary by rate-adjustability, points paid, property type, loan size and documentation (verification) level. Essentially, all of Chemical's loans are "salable", i.e., they are underwritten to secondary market guidelines. Both refinances and purchases of new homes are represented in the sample.

Once a loan is closed, it appears as a new observation on a panel system of loan performance that keeps track of all payments, prepayments, defaults and delinquencies. This system is used for servicing the mortgage (billing and collecting) so its reliability is high. The loan performance has been merged with all the applications data for this dataset.

Loans that prepaid before maturity could have refinanced, or the borrowers could have moved to a new residence. The two were separated out by searching over the New York City Department of Finance's filings of deeds, mortgages and uniform commercial codes (UCCs), using a database of public records owned by ***Public Data Corporation***.

---

[3]An "A" credit borrower has no credit filings (e.g., bankruptcies), no serious delinquencies in the past 2 years, few new credit inquiries in the past year, and is currently paid up on all accounts (Caplin, Freeman & Tracy, 1993).

This database is updated on a daily basis, and provides a very accurate way of finding out whether a homeowner has moved.

By searching over property addresses (block and lot number) we can see whether there was a mortgage filing and a deed transfer at the time the loan was prepaid. If there is a mortgage filing only, the prepayer must have refinanced the mortgage; if there was also a deed transfer, the prepayer must have moved. Movers who do not sell their property will not be identified.

For co-ops, the situation is more complicated because there is no deed for each individual unit. Owners are actually shareholders in the entire building, for which a single deed exists. Unlike condominium units, individual co-op units do not have their own lot numbers, but share one with all other units in the building. A transfer of shares in a co-op unit requires a UCC filing, but unfortunately so does a new mortgage against a co-op unit, so there is no easy separation of moves and refinances. The only way to distinguish between them is by matching the borrower and coborrower's names with that on the UCC filing. If a new UCC bearing the names of the borrowers was filed around the time of prepayment, then the loan is assumed to have refinanced. If the new filing does not bear the names of the borrowers, then they are assumed to have moved. A small error will arise from those who purchased another co-op in the same building.

5. Empirical Analysis of Mobility using Mortgage Applications Data

5.1 The sample

Two criteria are used for selecting our sample. First, because the sample begins in 1989, the break-even period for fixed rate mortgages (approximately 5 years) is not observed, while that for adjustable rate mortgages (ARMs) is much shorter and can be observed during the time frame of the data. Therefore, we restrict ourselves to ARM borrowers. This may not be representative of all homeowners because ARMs tend to be chosen by more mobile borrowers, as shown in ARM-FRM choice models (Brueckner & Follain, 1988).

Second, the sample is restricted to those loans that originated prior to 1992. Figure 5.1 shows that very few people are paying points after 1992, probably because borrowers had falling rate expectations. As shown in the previous section, the point choice does not reveal expected mobility in such an environment. The resulting sample consists of 1,421 observations, 157 that moved, and 1,264 that stayed. The behavior of these loans is observed until January 1994.

5.2 Splits by LTV

Figure 5.2 shows the distribution of loans by loan-to-value (LTV). The mass point at 80% LTV is explained by the fact that private mortgage insurance (PMI) has to be paid at LTVs of greater than 80%. The PMI is charged as a percentage of the entire loan balance and not just the marginal amount above 80% LTV, and so those who would have otherwise chosen an LTV slightly above 80% have a strong financial incentive to

Figure 5.1 Distribution of points

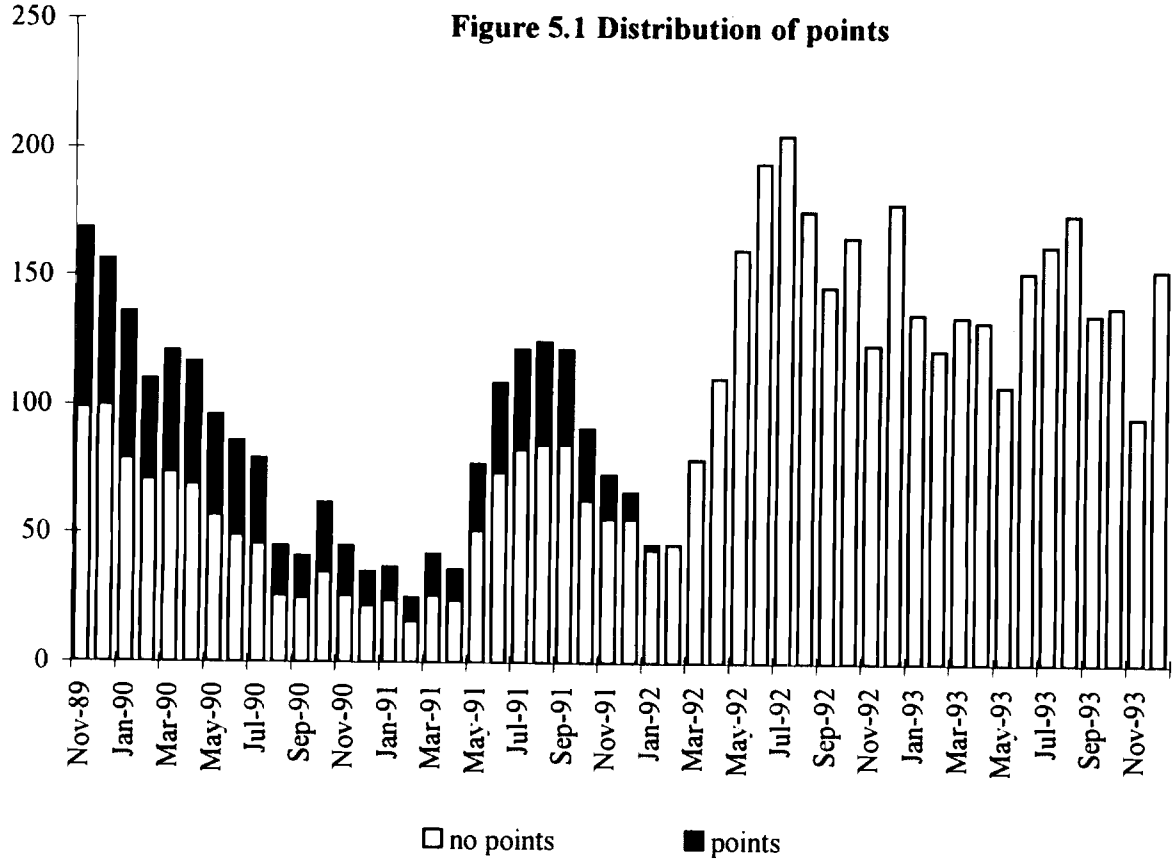
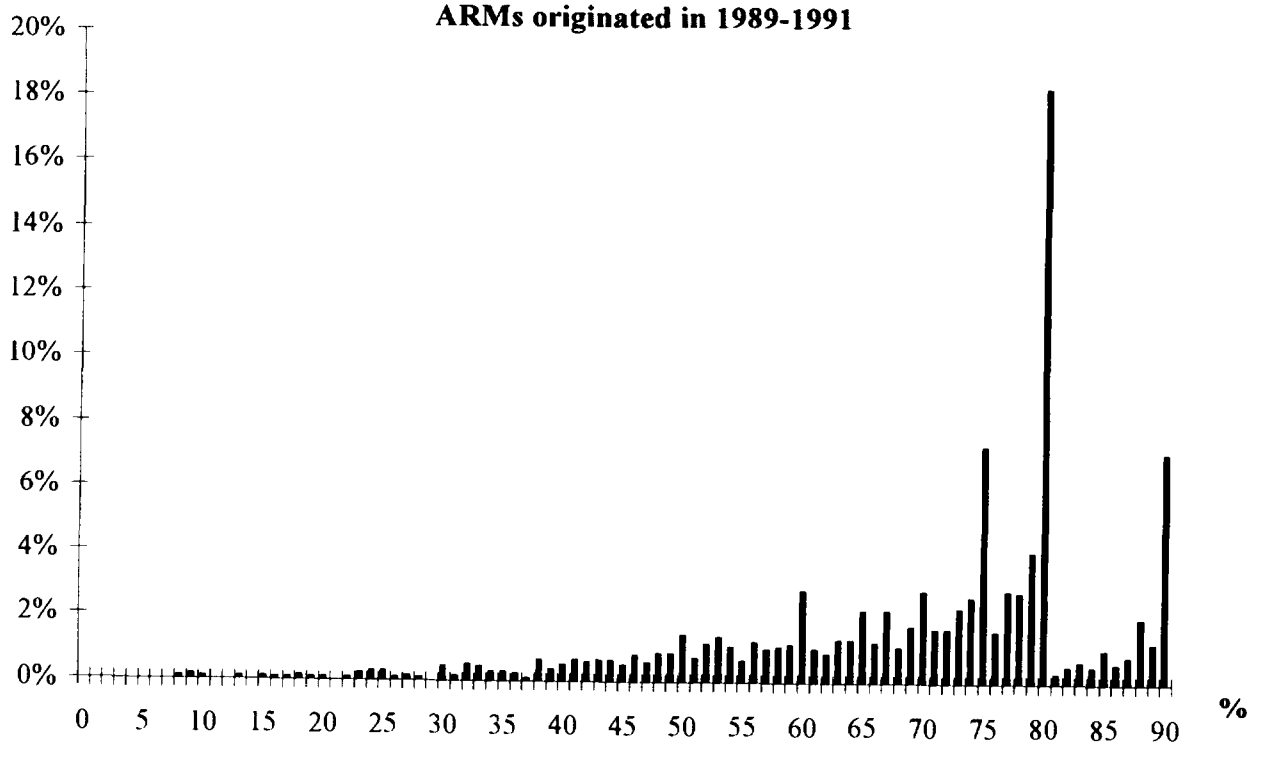


Figure 5.2: Distribution of Loans by LTV
ARMs originated in 1989-1991



accumulate a larger downpayment in order to reduce the ratio to 80%. Only those without the necessary upfront cash will pay the PMI. The mass point at 90% reflects the maximum LTV permitted by Chemical.

Table 5.1 shows that there are indeed substantial differences by LTV. Loans at or below the 80% cut-off are markedly different from those in the 80-90% band. The high LTV borrowers have lower incomes, even though loan amounts are similar to the low LTV group. High LTV borrowers are also much less likely to move. Given these differences, each group is analyzed separately.

5.3 Confirmation of the Expected Mobility Hypothesis

Summary statistics of the full sample by whether the household moved are provided in Table 5.2. On average, movers have higher appraised property values, lower LTVs and higher incomes. The demographic characteristics suggest that movers are generally at a later stage in their lifecycle: they are slightly older, more likely to be married, have more job tenure and are more likely to have been prior owners.

The movers are less likely to have paid points, consistent with the expected mobility hypothesis, and confirming Brueckner (1994)'s result. However, not surprisingly, borrowers do not have perfect foresight about their mobility decision and thus "mistakes" occur. Table 5.3 reports these events. The first column lists the percentage of those that did not pay points but did not move. Among the high LTV groups, this "mistake" is much more prevalent than in the low LTV groups.

Table 5.1: Sample Means by LTV (Medians for $ Values)

1989-1991 ARMs

	LTV ≤ 80%	**LTV > 80%**
moved (0,1)	0.12	0.06
loan amount $	104,200	113,900
appraised value $	160,000	130,000
loan to value %	66.42	87.91
initial interest rate %	9.11	9.55
paid points (0,1)	0.54	0.70
liquid assets $	12,600	12,600
total monthly income $	7,000	6,000
years at present type of work	12.2	10.4
years at present job	7.0	5.0
age	39.0	35.1
female (0,1)	0.20	0.19
married (0,1)	0.46	0.43
number of children	0.4	0.4
years of school	16.6	16.6
prior owner (0,1)	0.33	0.13
co-op (0,1)	0.46	0.38
number	1212	209

Demographic and employment variables refer to the borrower.
Income and assets are for the borrower and coborrower (if any) combined .

Table 5.2: Sample Means (Medians for $ Values)

1989-1991 ARMs

	Whole Sample	Non-movers	Movers
moved (0,1)	0.11		
loan amount $	106,400	102,100	135,200
appraised value $	158,000	150,000	205,000
loan to value %	69.58	70.25	64.15
initial interest rate %	9.17	9.16	9.28
paid points (0,1)	0.56	0.59	0.38
liquid assets $	12,600	12,600	11,700
total monthly income $	6,800	6,500	9,400
years at present type of work	11.9	11.8	13.0
years at present job	6.7	6.6	7.4
age	38.5	38.3	39.7
female (0,1)	0.20	0.21	0.17
married (0,1)	0.46	0.44	0.57
number of children	0.4	0.4	0.5
years of school	16.6	16.6	16.9
prior owner (0,1)	0.30	0.29	0.41
co-op (0,1)	0.44	0.49	0.35
number	1412	1264	157

Demographic and employment variables refer to the borrower.
Income and assets are for the borrower and coborrower (if any) combined .

For those borrowers who are at the maximum LTVs of 80% and 90%, cash constraints may have prevented them from paying points. Very little changes when these borrowers are ignored -- the "mistakes" are still more prevalent among the high LTV groups -- suggestive of equity constraints that affect high LTV loans more than low LTV loans. New York City experienced declines in house prices over this period, possibly locking in high initial LTV households who have negative equity and/or cannot afford a downpayment on a new home[4].

5.4 The Empirical Framework

The objective is to estimate mobility equations from the mortgage data and in particular, use expected mobility (in the form of the points indicator) as a predictor of actual mobility. Points become a sunk cost immediately after they are paid and do not enter the decision of whether to move ex post, after controlling for the interest rate difference. One can interpret the points indicator as simply a sorting device, similar to any other demographic dummy, which separates households into those who are likely to move and those who are not. Endogeneity problems would only arise if households had perfect foresight and as shown in Table 5.3, this is clearly not the case.

The model used is a version of the proportional hazard model developed by Flinn & Heckman (1983) that incorporates splines to fit the baseline hazard[5]. The hazard rate is the probability of moving after t periods conditional on not having moved before t periods. It is defined as:

---

[4]The median sales price of existing single family homes fell by 6.4% in nominal terms between 1989 and 1993 in the New York metropolitan region (New York State Association of Realtors, Inc).
[5]This methodology has been used by Gritz & MaCurdy (1992) and Cragg (1994), among others.

Table 5.3: Percentage of Sample that moved				
	No points		Points	
	Stay	Move	Stay	Move
whole sample	84.5	15.5	92.4	7.6
high LTV	95.2	4.8	93.9	6.1
high LTV without 90%	95.2	4.8	96.2	3.8
low LTV	83.3	16.7	92.0	8.0
low LTV without 80%	84.0	16.0	91.4	8.6

$$h(t, \alpha, \beta, X_t) = \exp [g(t, \alpha) + \beta X_t] \quad (1)$$

where t refers to duration in the house, X_t is a vector of time varying covariates and α and β are the parameters to be estimated.

The baseline hazard g is flexibly modeled as J sequential polynomials in time (the splines), with smooth transitions between each polynomial:

$$g(t,\alpha) = \sum_{j=1}^{J} w_j(t) \left[\alpha_{0j} + \alpha_{1j} t\right] \quad (2)$$

The parameters α determine the shape of each component and the weights w determine the amount of smoothness between the components. The weights are chosen to smooth the transition between two components so that the hazard rate is a weighted average of the two. The following smoothing factors are used:

$$w_j(t) = \Phi_j\left(\frac{\mu_j - t}{\sigma_j}\right) - \Phi_j\left(\frac{\mu_{j+1} - t}{\sigma_{j+1}}\right) \quad (3)$$

where Φ denotes the cumulative standard normal distribution function. The means locate the transitions while the standard deviations determine the amount of smoothness in the transitions.

5.5 The Low LTV Group

Only two polynomials were required to fit the data. The baseline function was specified as:

$$g(t, \alpha) = w_1(t) (\alpha_{00} + \alpha_{01} t) + w_2(t) (\alpha_{10} + \alpha_{11} t) \quad (4)$$

where

$$w_1(t) = 1 - \Phi\left(\frac{t-\mu}{\sigma}\right)$$
$$w_2(t) = \Phi\left(\frac{t-\mu}{\sigma}\right) \quad (5)$$

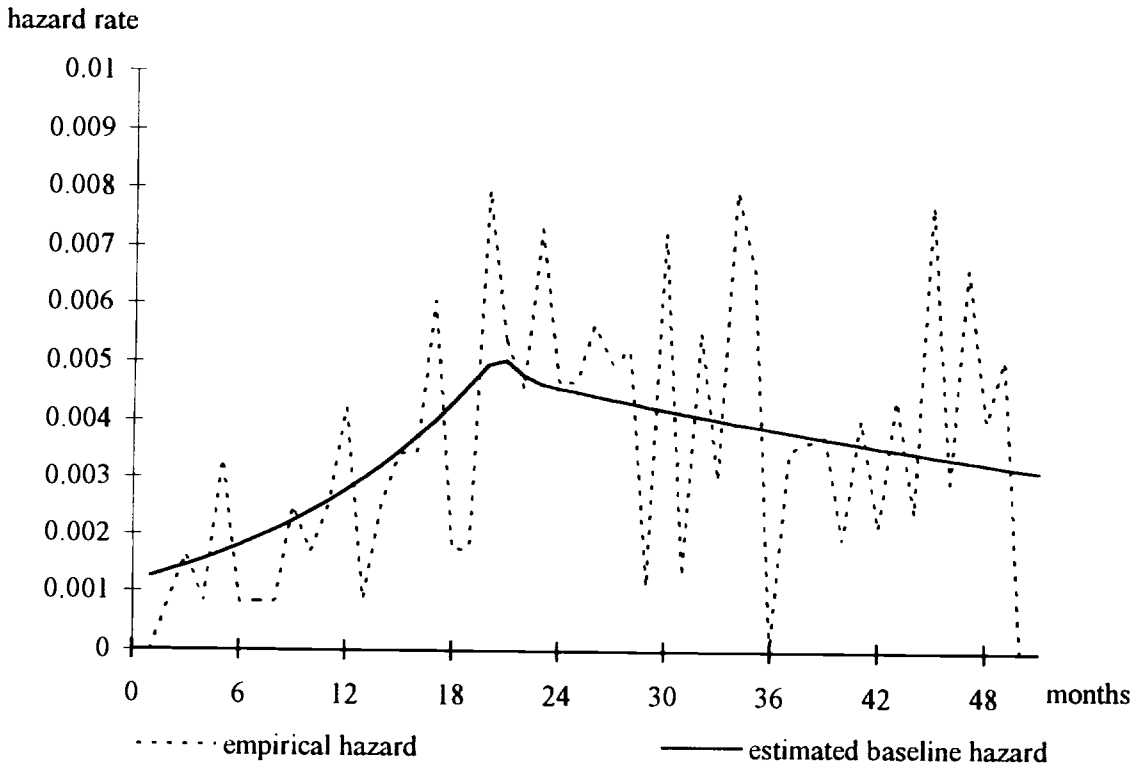
With extensive testing, the best fit was found to be one where $\mu = 31$ months and $\sigma = 0.75$. Figure 5.3 shows a plot of this against the empirical hazard.

Table 5.4 presents the results of putting other covariates in the hazard model. A larger coefficient implies that the probability of moving is higher as the covariate rises. Column 1 shows the pure effect of the interest rate incentive, the LTV and the loan amount. The interest rate incentive is defined as the difference between the current interest rate paid on the loan and the Freddie Mac ARM rate, both lagged two months[6]. As shown by Quigley (1987), among others, having a lower rate than market reduces the probability of moving because the interest rate advantage would have to be forgone. Similarly, having an interest rate that is higher than market increases the probability of moving because the opportunity cost of taking a new mortgage is reduced. This is true even for ARMs because of the interest rate caps and in some cases, the relatively long adjustment periods. The coefficient on interest rate incentive is significant and positive, in accordance with prior predictions.

The LTV ratio coefficient is negative and significant, consistent with the hypothesis that equity constraints are present. Borrowers with higher initial LTVs are more likely to have negative equity. The loan amount is interpreted as a proxy for wealth, and its coefficient is positive and significant. Again, this is consistent with equity

[6]The use of other interest rate series does not significantly affect the results because of high correlation. A two month lag is used since it would typically have taken this long to take out a new mortgage.

Figure 5.3: The baseline hazard for the Low LTV group



Estimated baseline hazard:

$$h(t) = -6.820 + 0.767\, w_{1t} + 1.886 w_2 - 0.018\, w_{2t}$$

Table 5.4: Proportional Hazard Estimates for the Low LTV Group

	1	2	3	4	5	6
points			-0.823		-0.752	-0.751
			(0.181)		(0.182)	(0.183)
rate incentive	0.313	0.329	0.284	0.327	0.306	0.317
	(0.073)	(0.074)	(0.075)	(0.074)	(0.075)	(0.076)
LTV	-2.611	-2.386	-2.570	-2.373	-2.297	-2.298
	(0.486)	(0.484)	(0.468)	(0.538)	(0.505)	(0.504)
loan amount	0.320	0.514	0.326	0.450	0.537	0.547
	(0.119)	(0.119)	(0.120)	(0.130)	(0.126)	(0.127)
income		-0.002				
		(0.028)				
liquid assets		0.014				
		(0.036)				
children				0.112		
				(0.098)		
school				0.012		
				(0.014)		
prior owner				0.169		
				(0.185)		
40 years				0.473	0.440	0.441
				(0.221)	(0.200)	(0.201)
55+ years				0.445	0.375	0.355
				(0.301)	(0.272)	(0.272)
female				-0.099		
				(0.237)		
married				0.303	0.369	0.367
				(0.194)	(0.179)	(0.179)
yrs type work				0.003		
				(0.013)		
years job				-0.006		
				(0.014)		
unemployment				0.018		
				(10.519)		
co-op						-0.238
						(0.187)
Log Likelihood	929.3	-924.7 [1]	-917.8 [2]	918.3	-909.8	-908.7

(Standard errors in parentheses)

[1] Likelihood ratio test (at the 5% level) against column 1 cannot reject that the additional coefficients are equal to zero.

[2] Likelihood ratio test (at the 5% level) against column 1 rejects that the additional coefficient are equal to zero.

constraints because wealthier borrowers can overcome the lock-in effects of negative equity with other assets.

Column 2 shows that income and liquid assets are insignificant in the presence of the loan balance variable. A joint likelihood ratio test shows that they do not add explanatory power.

Column 3 shows the effect of adding the points indicator. It significantly raises the likelihood function and is shown to be highly significant and of the expected sign. Paying points is associated with a lower hazard rate. Moreover, the other coefficients are largely unchanged.

Column 4 shows the effects of including other variables that predict mobility, excluding the points indicator. Being young or old increases the hazard rate of moving relative to the middle aged group. Being married and having children also increases the hazard. The employment variables do not have an effect, even if the other variables are removed from the estimation. The relative unemployment rate between New York City and the nation has a positive but insignificant effect.

Column 5 shows the points indicator with the demographic variables that worked in column 4. A likelihood ratio test rejects the restriction that the dummies on married, age <40 and age 55+ are equal to zero. The effect of the points indicator remains strong, showing that it provides additional explanatory power over the other demographic variables. Adding the co-op dummy in column 6 yields a negative, but not quite significant effect.

Figure 5.4: Estimated probability of moving for the Low LTV group

Typical borrower: 1 year ARM originated in January 1990,
Loan amount = $150,000, LTV = 75%, Age < 40 years, Married

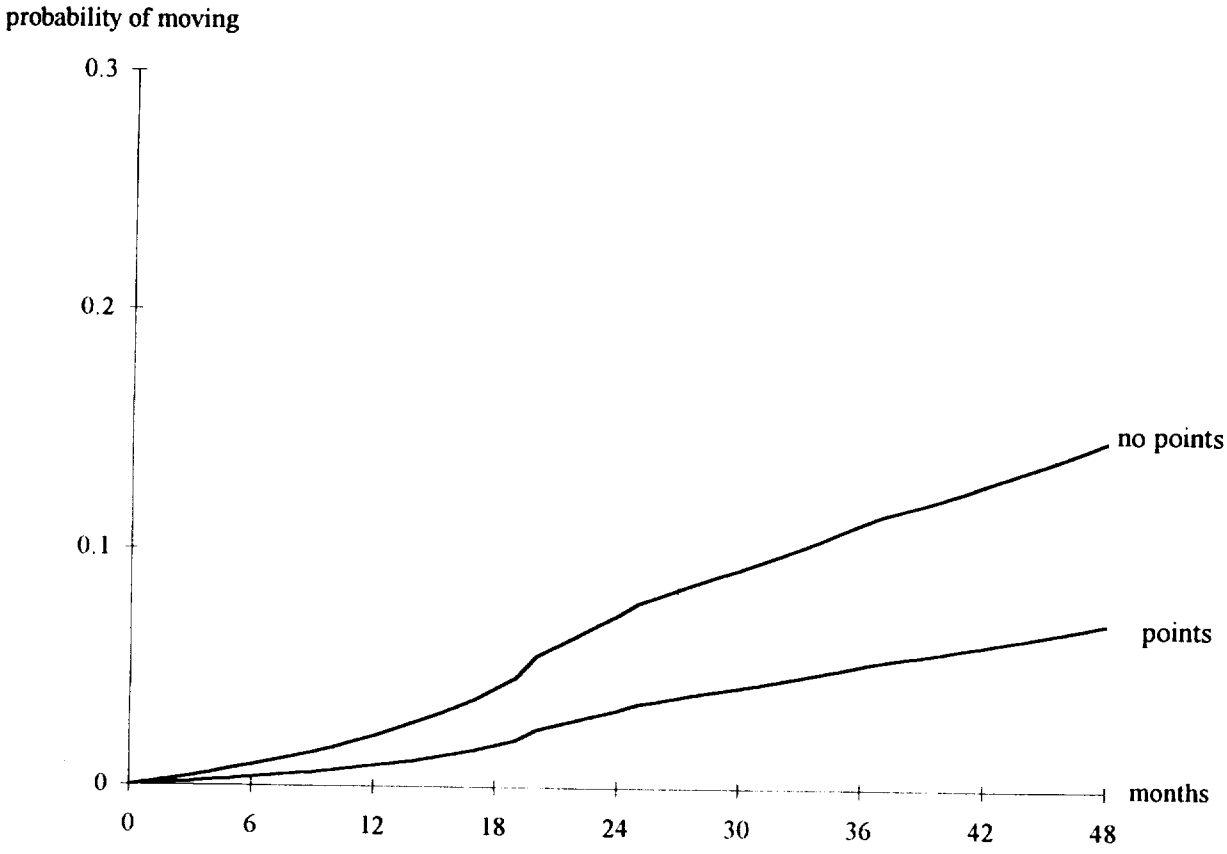


Figure 5.4 shows the estimated probability of moving for a "typical" borrower with and without points using the column 5 estimates. After two years, the moving probability associated with zero points is 7.3%, significantly higher than that associated with having paid points, 3.3%. After three years, this has increased to 11.1% compared with 5.2%.

5.6 The High LTV Group

Table 5.5 gives estimates for the high LTV group. The baseline hazard was estimated with one linear polynomial in time, and weights were not needed in this case. The covariates are all rather insignificant for this group, as shown in columns 1, 2 and 3. The rate incentive is barely significant. The coefficient on LTV does not provide explanatory power, possibly because there is very little variation for this group. The loan amount coefficient is insignificant and much smaller in magnitude than for the low LTV group.

Column 4 shows that points are not a good predictor of mobility: the coefficient is close to zero and insignificant. Borrowers are not moving as they had previously expected. Adding other demographic variables in column 5 does not improve the likelihood function.

There is reason to believe that some of the borrowers at 90% LTV may be liquidity constrained from paying any points. However, if these are removed from the analysis, the estimates do not substantially change.

The estimates are in stark contrast to the low LTV group and are very suggestive of large constraints to mobility affecting high LTV borrowers. The most likely culprit is

Table 5.5: Proportional Hazard Estimates for the High LTV Group

	1	2	3	4	5	6
points				0.006		0.006
				(0.876)		(0.894)
rate incentive	0.371	0.376	0.402	0.379	0.344	0.377
	(0.197)	(0.199)	(0.215)	(0.235)	(0.299)	(0.265)
LTV		-1.600				
		(10.836)				
loan amount	0.199	0.248	-0.149	0.363	0.307	0.360
	(0.570)	(0.583)	(0.849)	(0.637)	(1.130)	(0.649)
income			0.117			
			(1.086)			
liquid assets			0.368			
			(0.425)			
children					-0.185	
					(0.962)	
school					0.069	
					(0.169)	
prior owner					0.265	
					(1.338)	
<40 years					3.870	
					(48.561)	
55+ years					-0.720	
					(211.690)	
female					0.020	
					(1.068)	
married					0.491	
					(0.959)	
yrs type work					-0.057	
					(0.132)	
years job					-0.037	
					(0.300)	
unemployment					0.057	
					(31.062)	
co-op						-0.026
						(0.820)
Log Likelihood	-86.2	86.2 [1]	-85.5 [1]	-86.2 [1]	-80.6 [1]	-86.2 [1]

(Standard errors in parentheses)

[1] Likelihood ratio test (at the 5% level) against equation 1 cannot reject that the additional coefficients are equal to zero.

the decline in property values, coupled with downpayment requirements for the purchase of another home.

5.7 The Probabilities of Moving by LTV and Points

The coefficients from Tables 5.4 (column 5) and 5.5 (column 4) can be translated into probabilities of moving for different LTV classes and points. Figure 5.5 shows the probabilities for those that did not pay points. Moving increases with lower LTVs. Similarly for those that paid points in Figure 5.6. For high LTV loans, there is very little difference between those that paid and did not pay points: households are not moving as previously expected. The two graphs are drawn to the same scale and there is little difference between the 90% lines in both cases.

6. Conclusions

Mortgage applications are a detailed and accurate source of demographic, income and asset information that is verified by underwriters. This paper has discussed how the accuracy and detail of mortgage data can be applied to important areas of economics outside of mortgage finance, in particular, mobility and location choice.

As a supplement to the variables from the application form, the self-selection of mortgage contracts has been used to infer expected mobility from the choice of points. The points indicator was tested, using the Chemical Bank dataset of mortgage loans, in a

Figure 5.5: Probability of Moving for the No Points Group, by LTV

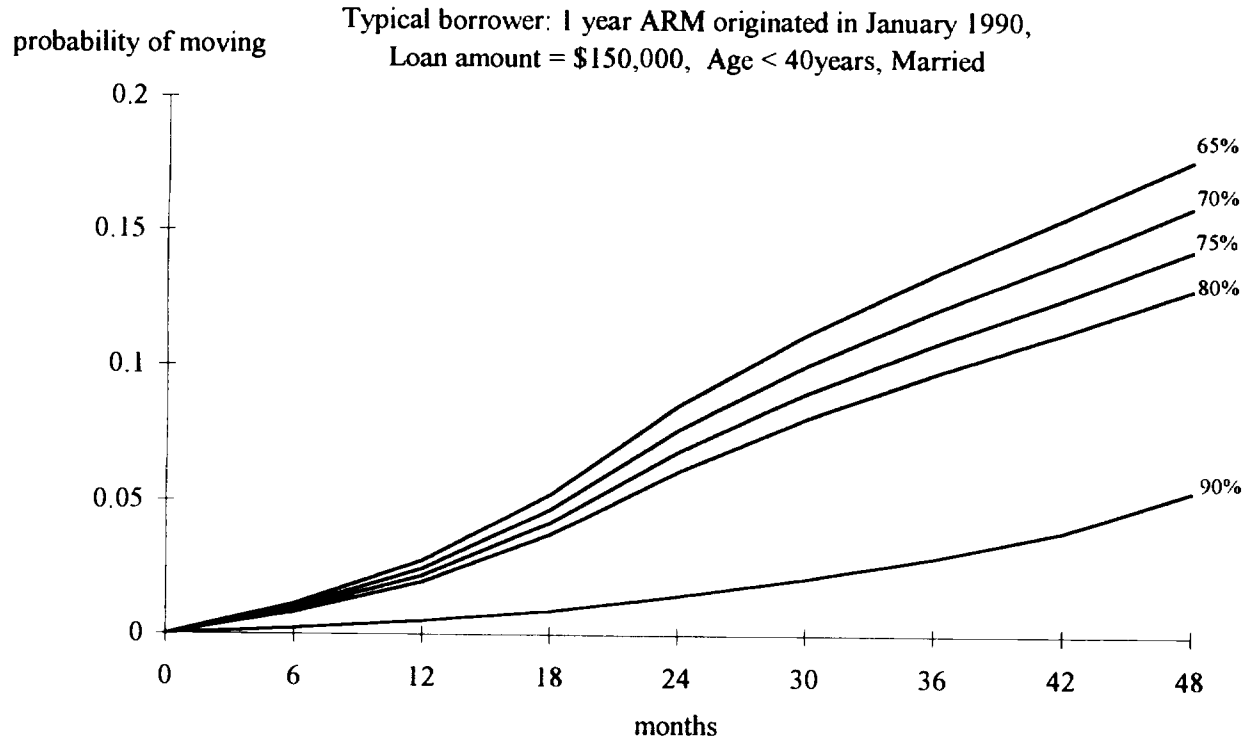
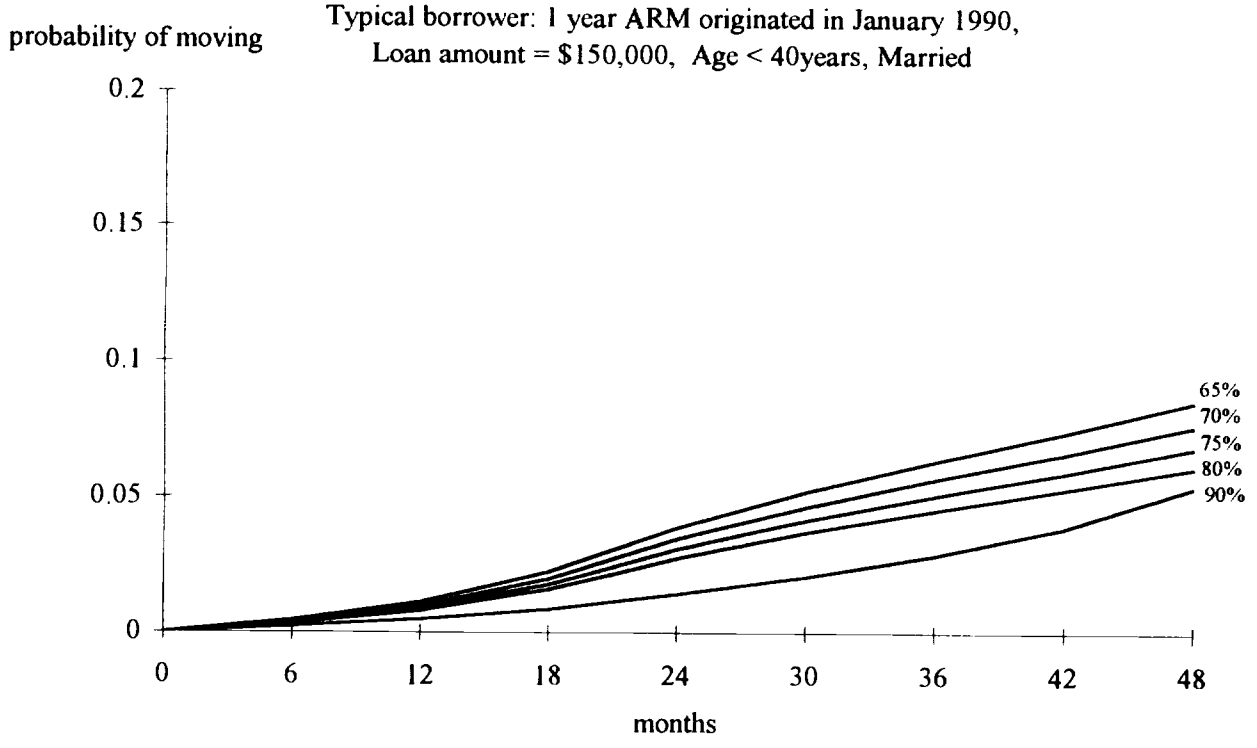


Figure 5.6: Probability of Moving for the Points Group, by LTV



duration model of housing spells. The points indicator was shown to be highly significant in predicting mobility for the low LTV group and provided explanatory power beyond the other variables. The mobility of the high LTV group could not be explained by this indicator, nor by any of the other standard demographic and income variables that we believe to influence mobility.

The substantial difference for high and low LTV borrowers is evidence for the presence of large constraints to mobility for the high LTV group that have arisen from the recent decline in property values and the downpayment requirements for the purchase of another home. Future work will extend the sample geographically so that variation in house price movements can be included in the analysis.

These findings demonstrate the potential fruitfulness of using mortgage applications data. Its accuracy and wide applicability more than adequately reward the efforts required in collecting data of this kind.

References

Bartel, Ann P., 1979, The Migration Decision: What Role Does Job Mobility Play?, American Economic Review.

Blanchard, Olivier and Lawrence Katz, 1992, Regional Evolutions, Brookings Papers on Economic Activity.

Borjas, George J., Stephen G. Bronars and Stephen J. Trejo, 1992, Self-Selection and Internal Migration in the United States, Journal of Urban Economics.

Brueckner, Jan K. and James R. Follain, 1988, The Rise and Fall of the ARM: An Econometric Analysis of Mortgage Choice, Review of Economics and Statistics .

Brueckner, Jan K., 1994, Borrower Mobility, Adverse Selection and Mortgage Points, Journal of Financial Intermediation.

Caplin, Andrew, Charles Freeman and Joseph Tracy, 1993, Collateral Damage: How Refinancing Constraints Exacerbate Regional Recessions, NBER Working Paper No. 4531.

Cragg, Michael, 1994, The Dynamics of Welfare Participation in Canada, Columbia University Working Paper.

Flinn, Chris and James Heckman, 1983, The Likelihood Function for the Multistate-Multiepisode Models in Models for the Analysis of Labor Dynamics, in: eds., R. Bassmann and G. Rhodes, Advances in Econometrics, vol. 3.

Genesove, David and Christopher J. Mayer, 1994, Equity and Time to Sale in the Real Estate Market, NBER Working Paper No. 4861.

Gritz, Mark and Thomas MaCurdy, 1992, Transitions from Welfare to Work, Report prepared for the U.S. Department of Health and Human Services.

Gronberg, Timothy J. and W. Robert Reed, 1992, Estimation of Duration Models Using the Annual Housing Survey, Journal of Urban Economics.

Harmon, Oskar R. and Michael J. Potepan, 1988, Housing Adjustment Costs: Their Impact on Mobility and Housing Demand Elasticities, Journal of the American Real Estate and Urban Economics Association.

Henderson, J. Vernon and Yannis M. Ioannides, 1989 Dynamic Aspects of Consumer Decisions in Housing Markets, Journal of Urban Economics.

Kiel, Katherine A., 1994, The Impact of House Price Appreciation on Household Mobility, Journal of Housing Economics.

Lichter , Daniel T, and Gordon F. De Jong, 1990, The United States, in: eds., Charles B. Nam and William J. Serow, The International Handbook on Internal Migration.

Nakagami, Yasuhiro and Alfredo M. Pereira, 1991, Housing Appreciation, Mortgage Interest Rates, and Homeowner Mobility, Journal of Urban Economics.

Potepan, Michael J., 1989, Interest Rates, Income, and Home Improvement Decisions, Journal of Urban Economics.

Quigley, John M., 1987, Interest Rate Variations, Mortgage Prepayments and Household Mobility, Review of Economics and Statistics.

Stein, Jeremy C., 1993, Prices and Trading Volume in the Housing Market: A Model with Downpayment Effects, NBER Working Paper No. 4373.

Stone, Charles A. and Anne-Marie Zissu, 1990, Choosing a Discount Point/Contract Rate Combination, Journal of Real Estate Finance and Economics.

Tiebout, Charles M., 1956, A Pure Theory of Local Expenditures, Journal of Political Economy.

Yang, T.L. Tyler , 1992, Self-Selection in the Fixed Rate Mortgage Market, Journal of the American Real Estate and Urban Economics Association.