Markov Chain Model of Rent Burden in the Housing Choice Voucher Program

Brent Mast
U.S. Department of Housing and Urban Development

Introduction

Since 1937, the United States has made access to affordable housing a national priority. For a variety of reasons, assistance has been provided through housing subsidies rather than general income transfers.

Affordable housing assistance was initially provided through public housing projects run by local Public Housing Agencies (PHAs). In 1965, PHAs began experimenting with providing opportunities in privately owned rental units.[1] Private opportunities have been provided through a variety of programs.

There are numerous arguments for providing assistance in privately owned buildings instead of public housing. The primary motivation for increasing private sector housing choices has been expanding social and economic opportunities for low-income households receiving housing assistance. Another argument is that private owners might have better incentives for operational efficiency, thus lowering program costs.

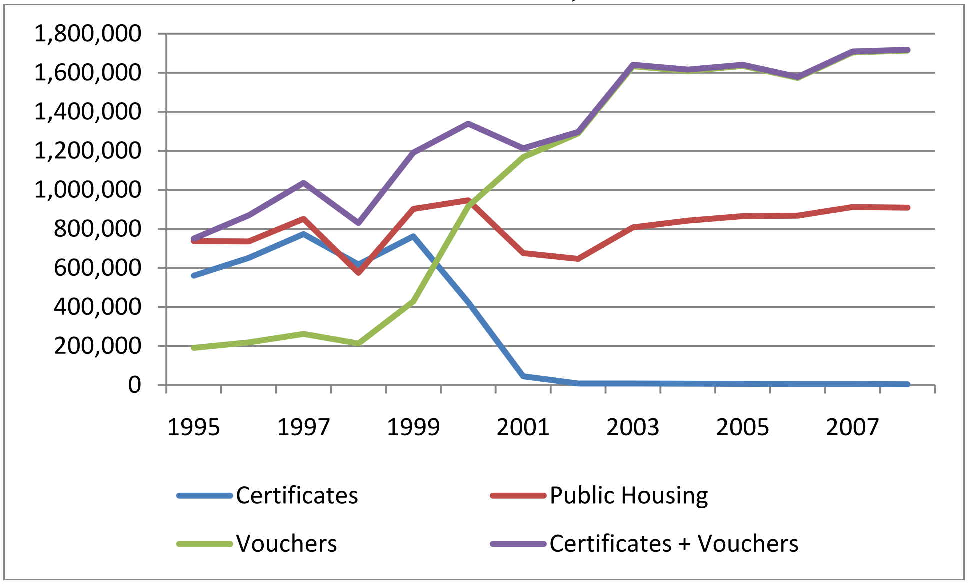
Exhibit 1 reports trends in rental assistance for 1999-2008 through public housing and the two largest private options – Certificates and the Housing Choice Voucher Program (HCVP). In 1995 there were .74 million households in public housing, compared to .56 million certificates and .19 million vouchers.

Certificates have been phased out over time, while Housing Choice Vouchers have increased dramatically. Today, HUD's Section 8 program provides assistance to about .9 million households in public housing projects, and 1.7 million through the Housing Choice Voucher Program. HCVP is the now largest U.S. rental assistance program, with an annual budget of over $16 billion.

Voucher recipients choose their own rental units, and rent is paid for partially by the household and partially through HUD payments to landlords. The voucher program is administered by approximately 2600 local Public Housing Agencies (PHAs). Among the PHAs' responsibilities are selecting applicants to receive vouchers, annual verification of household income, and determination of each household's rent responsibility.

[1] See http://www.mphaonline.org/section8.cfm for a concise history of HUD rental assistance programs.

Exhibit 1: Trends in HUD Rental Assistance, 1995-2008



The Housing Assistance Payment (HAP) under the program is the difference between 1) the minimum of gross rent (rent + utility costs) and the Payment Standard, and 2) the Total Tenant Payment (TTP). Payment Standards and TTP are both set by PHAs. TTP is the minimum household contribution, which is usually about 30 percent of household adjusted income.

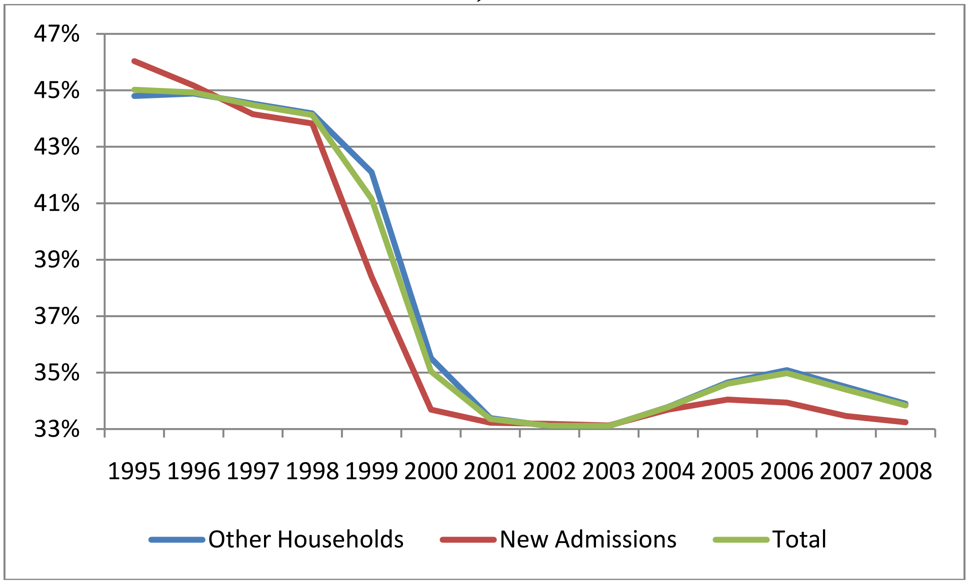
If the Payment Standard set by the PHA is in-line with actual rents, households should be able to find units meeting HUD quality standards at burden levels around 30 %. Yet in reality, many households face burdens considerably greater than 30 %.

Rent burdens above 30 percent don't necessarily imply inadequate subsidies. The program is designed to foster choice. Households may choose burden levels above 30 % for a variety of reasons. For instance, households may have a preference for larger homes, newer homes with better amenities, or homes in better neighborhoods.

Starting in 1998, when admitted to the program or moving into a new unit, a family's rent burden is required to be between 30 and 40 % of adjusted household income. Data indicate that in practice, however, households admitted to the program or moving into new units often have rent burdens in excess of 40 %. In other years, rent burden has a 30 % floor but no ceiling.

Exhibit 2 reports trends in rent burden for new admissions, other households, and total households from 1995-2008. After restrictions were placed on burdens at admission in 1998, average rent burdens fell dramatically. Around 2003 burdens increased, and fell again starting around 2006. As of 2008, the mean burden was 33.837 % overall, 33.244 % for new admissions, and 33.896 % for prior admissions.

Exhibit 2: Mean Rent Burden Trends, 1995-2008



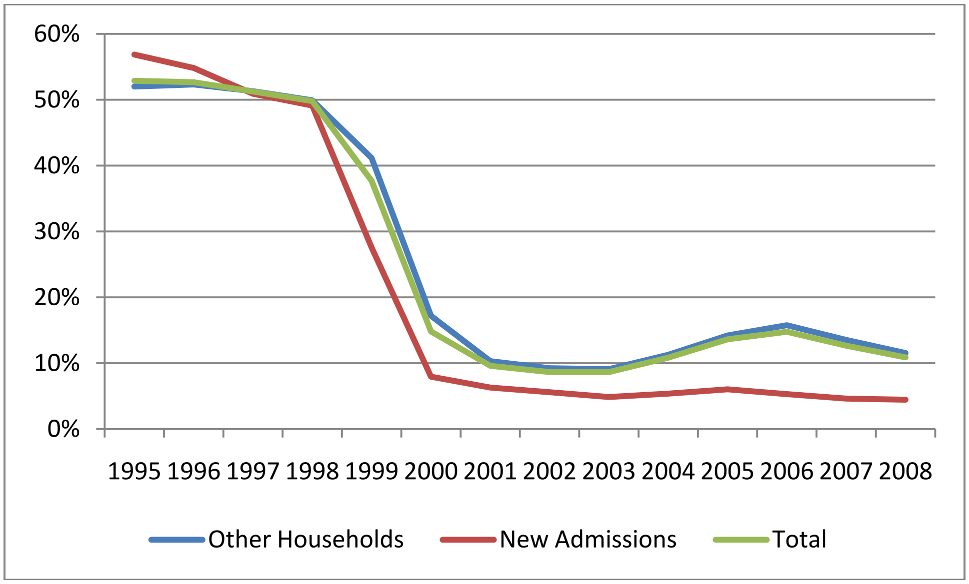
While mean burdens in recent years may seem reasonable, many households have burdens considerably higher than the mean. Exhibit 3 reports trends for percentages of households with burdens of at least 41 %. Since 2000, the fraction of households with burdens in excess of 40 % has fluctuated. In 2008, 10.893 % of household fell in this category. The fraction of new admissions with burdens of at least 40 % shows less volatility, equaling 4.464 % in 2008.

Policy implications depend not only on the number of households with high rent burdens at any given time, but also on the tendency of households to have persistent periods of high burden. While prior research suggests considerable turnover in the population of households with high rent burdens (HUD 2005), prolonged burden may disproportionately affect the most disadvantaged households (Susin 2007).

This study models the likelihood of persistent high rent burden in the HCVP as a Markov Chain. Longitudinal data are analyzed on households admitted to the program between 2000 and 2008. Estimates imply increasing burden with time in the program and considerable household mobility.

The next section summarizes relevant literature. Program guidelines for determining rent burden are then discussed. The data are described next. Markov Chain analysis follows. Conclusions are summarized in the final section.

Exhibit 3: % of Households with Burdens >=41 %



II Literature Review

Numerous studies have examined rent burden in the general population. For instance, Susin (2007) analyzes duration of rent burden for participants in the 2001 Survey of Income of Program Participation (SIPP). The focus of the study was identifying household characteristics related to prolonged periods of high burden. He reports numerous indicators of need, such as non-employment and receipt of Supplemental Security Income (SSI), are associated with persistent high burden.

A few studies have also examined rent burden in the HCVP program. McClure (2005) is the most recent, using the same data source as this study (HUD's MTCS data system, described below in section IV). Analyzing household data for 2000 to 2002, he studies how rent burden varies with a large number of household, housing market, and policy variables. He finds high rent burdens are most prevalent among very low income families.

Using only three years of data, McClure's study doesn't indicate whether periods of extremely low income (and high rent burden) are "chronic or transitory" (McClure 2007, p. 18).

Susin (2007) uses regression analysis to predict duration of high rent burden. McClure (2007) only reports summary statistics. I am not aware of any studies analyzing rent burden as a Markov Chain.

Many studies have applied Markov Chain models to other housing market topics, however. An early example is Clark (1965), who studies how Markov Chains can be used to predict geographic movement of households.

III Rent Burden Formula

Rent burden (burden) equals the difference between gross rent including utility payments (rent) and the Housing Assistance Payment (HAP), divided by household adjusted income (income):

$$burden = [rent - HAP]/income \quad \text{(eq1)}$$

HAP is the difference between 1) the minimum of gross rent and the Payment Standard (PS), and 2) the Total Tenant Payment (TTP):

$$HAP = \min(rent, PS) - TTP \quad \text{(eq2)}$$

TTP is the minimum household contribution. It is calculated as the maximum of four values: 1) 30 percent of household adjusted income, 2) 10 percent of gross income, 3) the welfare rent, and 4) the PHA minimum rent. For the vast majority of households, TTP is 30 percent of household adjusted income. Under this assumption, eq2 can be rewritten as:

$$HAP = \min(rent, PS) - .3*income \quad \text{(eq3)}$$

Note that until 1988, households could keep the difference between Payment Standards and rent when rent fell below Payment Standards. Thus households had an incentive to search for units renting below the Payment Standard in order to reduce their rent burdens.

Now, however, the minimum rent burden a household can incur is at least 30 percent. Thus households have no incentive to search for qualifying units renting below the Payment Standard in order to reduce their rent burden.

Yet households may select units with gross rent below the payment standard for a variety of other reasons. Indeed, a slight majority of households fall in this category. Accordingly, let H1 equal the maximum of zero and Payment Standard minus gross rent:

$$H1 = \max(0, PS - rent) \quad \text{(eq4)}$$

Thus eq3 can be rewritten as follows, where H1 is nonnegative:

$$HAP = PS - H1 - .3*income \quad \text{(eq5)}$$

Let r equal the ratio of the Payment Standard to Fair Market Rent (FMR):

$$r = \text{Payment Standard}/FMR \quad \text{(eq6)}$$

The Payment Standard equals r multiplied by FMR:

$$PS = r*FMR \quad \text{(eq7)}$$

Substituting eq7 into eq5:

$$HAP = r*FMR - H1 - .3*income \quad \text{(eq8)}$$

Substituting eq8 into eq1:

$$burden = .3 + H1 + rent/income - r*FMR/income \quad \text{(eq9)}$$

Actual rent at the 40th or 50th percentile of the rent distribution in the FMR region for a given number of bedrooms (FMR*) equals HUD's measure (FMR) plus measurement error (error):

$$FMR^* = FMR + error \quad \text{(eq10)}$$

Gross rent is actual rent at the 40th or 50th percentile of the rent distribution in the FMR region (FMR*) plus a household preference factor, H2, for a unit with gross rent above or below FMR*:

$$rent = H2 + FMR^* \quad \text{(eq11)}$$

Substituting eq10 into eq11:

$$rent = H2 + FMR + error \quad \text{(eq12)}$$

Substituting eq12 into eq9:

$$Burden = .3 + H1/income + H2/income + FMR/income - r*FMR/income + error/income \quad \text{(eq13)}$$

Simplifying:

$$Burden = .3 + (H1 + H2)/income + FMR(1-r)/income + error/income \quad \text{(eq14)}$$

where H1>=0, and (H1 + H2)/income + FMR(1-r)/income + error/income is nonnegative as well. That is, rent burden can't fall below 30 percent.

So rent burden can modeled as function of three main factors:

A) Household housing preferences. In particular, tendencies to select units with gross rent 1) below the Payment Standard, and/or 2) above or below rent at the actual 40th or 50th percentile of the rent distribution. Note that the incentive to select units renting below the Payment Standard diminished after 1988 when doing so no longer reduced rent burden.

B) Payment Standard deviations from FMR interacted with FMR, relative to household income; and

C) Deviations in FMR from actual 40th or 50th percentile rents, relative to household income.

IV. The Data

Data are extracted from HUD's MTCS/PIC data system. The system has quarterly entries for each family receiving Section 8 assistance starting in 1995. Data are available on rent burden and a large number of other household characteristics.

For the remainder of the study, I explore data on families admitted to the HCVP program between 2000 and 2008. Families admitted before 2000 are excluded due to the rapid decline in rent burden from 1995-1999 (see Exhibit 2).

I exclude Certificate households because 1) their rent burden regulations differ from HCVP regulations, and 2) there are few Certificates today, so the program is not of great policy importance. I also exclude HCVP households in the Homeownership program and HCVP households that were homeless prior to admission.

I also drop some outliers with suspect data. I exclude households if: 1) adjusted annual income is below $250 or above $40,000; or 2) gross rent (including utilities) is above $3,000 a month; or 3) rent burden as a fraction of adjusted or income is less than 29 % or more than 90 %.

The system is action based. The most common actions are 1) admissions; 2) annual re-exams; 3) interim re-exams due to changes in eligibility factors such as income or family size; 4) moves; and 5) exits from the program. The system captures the most recent action at end of each quarter. If multiple actions for a household occur during a quarter, only the most recent is available. If there is no action during a quarter, the family's entry is a duplicate of the entry for the previous quarter.

Rent contracts are effective for one year, and most households have only one action per year. Therefore most changes in rent burden are annual (not quarterly). Accordingly, for this study I employ longitudinal research files that capture the most recent action at the end of each year for each family. Compared to the quarterly data, the longitudinal data have been edited somewhat to make them more comparable across years. For instance, variable names that change over time are standardized, and data items from multiple tables are combined into a single file.

V Summary Statistics

Exhibit 4 reports household counts by year. Counts are reported for new admissions, households admitted in prior years, and all households. For 2000, there are 186,855 families, all of which were admitted the same year. In 2008 there are 638,989 households and 136,970 new admissions.

Exhibit 4: Household Counts by Year

Year	New Admissions	Other Households	All Households
2000	186855	0	186855
2001	178820	90325	269145
2002	186275	179024	365299
2003	155341	314332	469673
2004	112431	375327	487758
2005	106887	407502	514389
2006	136377	403519	539896
2007	147502	458796	606298
2008	136970	502019	638989

Total	1347458	2730844	4078302

In total, there are 4,078,302 observations on 1,347,458 HCVP families. Exhibit 5 depicts the distribution of observations per family. The median number of observations per household is 2, and the mean is 3.027. Only 5 % of families have 7 or more observations.

Exhibit 5: Observations per Household

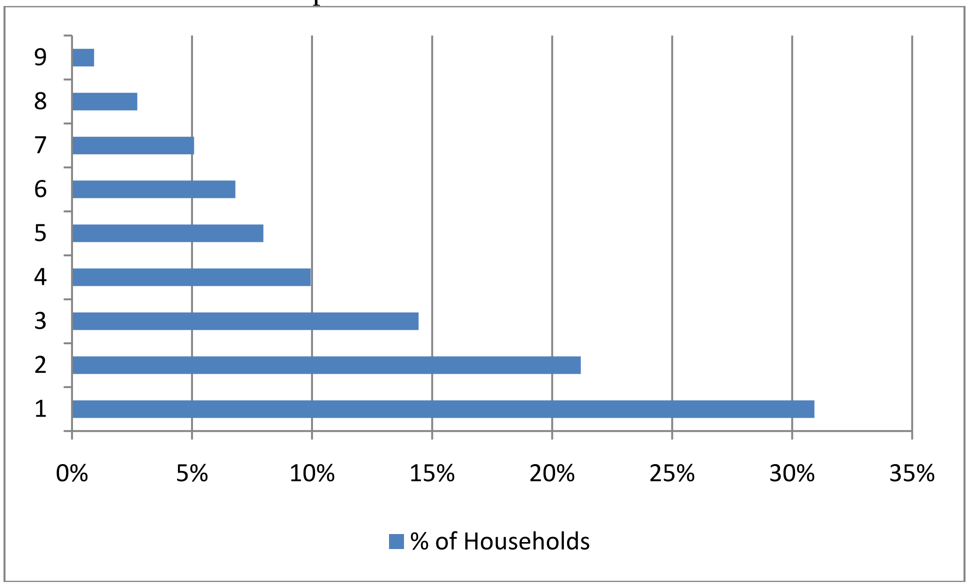


Exhibit 6 reports mean burden by years in the program. Exhibit 7 reports trends in percentages of households with rent burdens in excess of 40 %. Trends are reported for four household categories: 1) households with an elderly head or spouse (16.07 % of sample); 2) non-elderly households where the head or spouse has a disability (29.00 % of sample); 3) other households with children (45.77 % of sample); and 4) other households without children (9.16 % of sample).

Exhibit 6: Mean Burden by Tenure

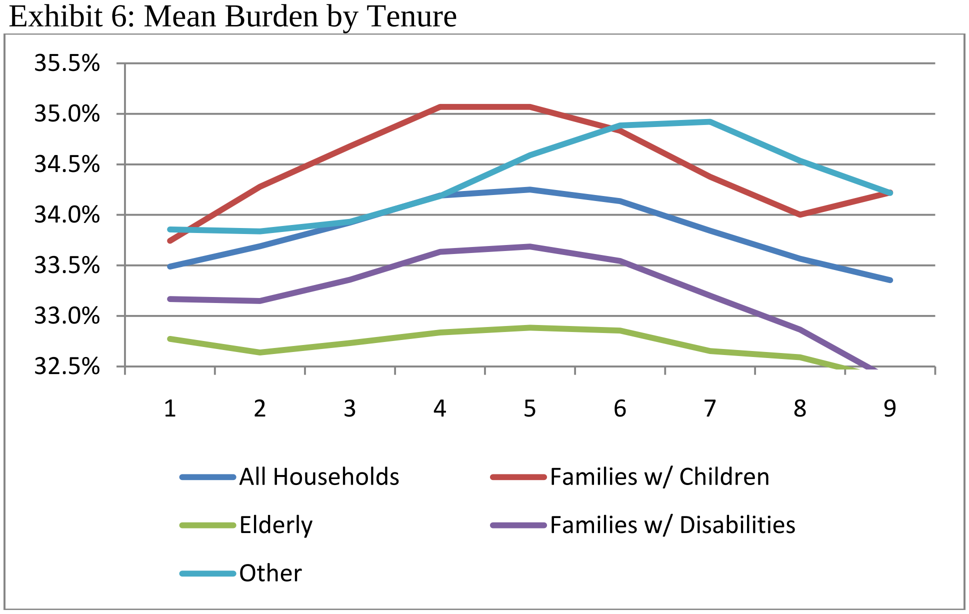
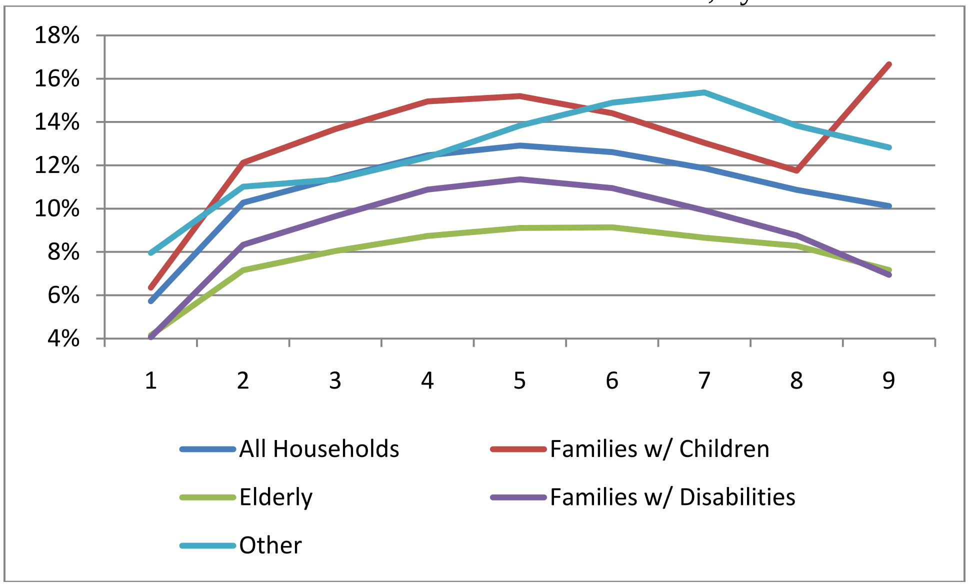


Exhibit 7: Percent of Households with Burdens >=41%, by Tenure



The patterns of mean burdens and high burdens are similar for all household categories. At admittance, burdens are relatively low. This is not surprising, given that initially burdens are not supposed to exceed 40 %. Over time, burdens increase and start to level off in the 4th or 5th year. While burdens start to decrease around year 5, this may be an artifact of rapidly dwindling observations over time (see Exhibit 5).

VI Markov Chain Model

Data

In this section, I model HCVP rent burden as a discrete time stochastic process. For brevity, I limit my analysis to households with children at admission. While household status can change over time, this is just one factor among many possibly affecting changes in rent burden.

Exhibit 8 reports households analyzed by year. Counts are reported for new admissions, prior admissions, and all households. In 2000 there are 32,680 families, all of which are new admissions. By 2008, there are 280,892 households and 66,649 new admissions. In total, the Markov Chain model is based on 1,737,608 observations on 629,093 families.

Exhibit 8: Household Counts, Families w/Children

Year	New Admissions	Other Households	All Households
2000	32680	0	32680
2001	98310	2078	100388
2002	100199	61613	161812
2003	81093	130327	211420
2004	58338	161149	219487
2005	52016	175202	227218
2006	67425	169929	237354
2007	72383	193974	266357

2008	66649	214243	280892
Total	629093	1108515	1737608

Assumptions

I assume a household's burden falls into four possible states:

1) Normal burden -- rent burden <=31 % ;
2) Medium burden -- 32 %<=rent burden<=40 %;
3) High burden -- 41 %<=rent burden<=49 %; and
4) Extreme burden -- rent burden>=50 %.

Burden at admittance is known. Burden in subsequent years satisfies the following assumptions:

1) When the household is in state i, there is a fixed probability P_{ij} of being in state j in the next year;

2) The conditional distribution of any future state X_{n+1} is independent of past states, depending only on the current state X_n.

Under these assumptions, rent burden is a Markov Chain.

Transition Probabilities

Exhibit 9 reports transition probability matrices P^n for years n=1 to 4. They were calculated by cross tabulating household states in year 0 (the family's admission year) with their states in years 1 through 4.

While up to 9 years of data are available, only a small fraction of households have more than 5 observations (see Exhibit 5). Survival analysis indicates the median time in the program is around 3 years, and only 25 % of households stay longer than 7 years (Cortes et al. 2008).

Suppose a family is in state 1 (normal burden) during year 0. P^1 indicates a .811 chance of remaining in state 1 for year 1. The chance of medium burden is .131, the chance of high burden is .027, and the chance of extreme burden is .030.

Exhibit 9: Transition Probability Matrices

P^1

0.811	0.131	0.027	0.030
0.358	0.459	0.105	0.077
0.326	0.303	0.221	0.150
0.344	0.196	0.122	0.338

P^2

0.723	0.185	0.046	0.046
0.441	0.371	0.102	0.087
0.419	0.300	0.146	0.135
0.422	0.232	0.105	0.241

P^3

0.664	0.219	0.059	0.057
0.456	0.349	0.102	0.092
0.459	0.287	0.127	0.127
0.450	0.251	0.099	0.200

P^4

0.643	0.232	0.064	0.061
0.478	0.330	0.102	0.090
0.476	0.287	0.118	0.120
0.489	0.249	0.090	0.172

Recall that rent burdens tend to increase in years 1 through 4 (see Exhibits 6 and 7). For a family with normal burden in year 0, their chance of normal burden decreases monotonically with program tenure. By year 4, their probability of normal burden is .643, while their chance of high or extreme burden has increased to .064+.061=.125.

While there is a trend toward increasing burden, there is also considerable mobility across states over time. For a family with extreme burden at admission, their chance of normal or medium burden in year 1 is .344 + .196=.540. Their chance of normal or medium burden in year 2 is .422 + .232=.654. In year 3 the chance increases to .450 + .251=.701. By year 4 the probability of normal or medium burden is .489 + .249=.738.

Let excess burden refer to state 3 or 4, with burden>=.41. Suppose a family has extreme burden (state 4) at admission. What's their probability of 2 periods of excess burden in years 1 through 4? There are 6 possible ways this could happen: S={B_1B_2 , B_1B_3, B_1B_4, B_2B_3, B_2B_4, B_3B_4}, where B_i=excess burden in year i. Redefine the sample space S as {E_1, E_2, E_3, E_4, E_5, E_6}. The probability of 2 periods of excess burden equals P(E_1 or E_2 or E_3 or E_4 or E_5 or E_6). This equals $\sum_i P(Ei)$ – $\sum_{i<j} P(EiEj)$ + $\sum_{i<j<k} P(EiEjEk)$ – $\sum_{i<j<k<l} P(EiEjEkEl)$ + $\sum_{i<j<k<l<m} P(EiEjEkElEm)$ - P($E_1E_2E_3E_4E_5E_6$). In words,

"the probability of the union of n events equals the sum of the probabilities … taken one at time … minus the sum of the probabilities … taken two at a time plus the sum of the probabilities … taken three at a time, and so on"(Ross 2007, p9).

Exhibit 10 reports probabilities of 2 periods of excess and extreme burden for families with states 1 through 4 at admission. Starting with extreme burden, over the next 4 years there is a 52.029 % chance of 2 periods of excess burden, and a 28.951 % chance of 2 years of extreme burden. The probabilities decrease with burden at admission. For families admitted with normal burden, the chance of 2 years of excess burden is 5.478 %, and the chance of 2 years of extreme burden is 1.387 %.

Exhibit 10: Probability of 2 Years of Excess and Extreme Burden, Years 1-4.

Admission State	Prob of 2 years of excess burden (burden>=40 %)	Prob of 2 years of extreme burden (burden>=50 %)
1: Normal burden <=32 %	5.478%	1.387%
2: Medium burden 32-40 %	19.671%	4.396%
3: High burden 41-49 %	39.722%	10.109%
4: Extreme burden >=50 %	52.029%	28.951%

Predictions
The transition probability matrices are based on data for all years. Yet the probabilities might change over time, leading to worse predictions. Exhibit 11 reports predicted versus actual year 4 outcomes for households admitted in 2000 and 2004. Predicted outcomes equal X_0P^4, where X_0 is a row vector of household shares for year 0 (admission), and P^4 is the 4th year transition probability matrix (see Exhibit 8). In 2000, X_0={.628, .284, .044, .044}. In 2004, X_0={.572, .367, .028, .032}.

Exhibit 11: Year 4 Actual and Predicted Outcomes, 2000 and 2004 Admissions

2000 Admissions

Burden	Actual	Predicted	Difference
30% -31%	55.118%	58.162%	-3.045%
32% - 40%	27.826%	26.298%	1.527%
41% - 49%	8.652%	7.871%	0.781%
50% &Above	8.404%	7.668%	0.736%

2004 Admissions

Burden	Actual	Predicted	Difference
30% -31%	65.655%	57.242%	8.413%
32% - 40%	23.328%	27.008%	-3.680%
41% - 49%	5.978%	8.071%	-2.093%
50% &Above	5.039%	7.680%	-2.640%

Predictions are closer for 2000 admissions compared to 2004. For example, 8.652 % of 2000 admissions had high burdens (41-49 %) in 2004. The model predicted 7.871 %, a difference of less than 1 %. 5.978 % of 2004 admissions had high burdens in 2008. The model predicted 8.071 %, a difference of 2.093 %. The differences between years are not huge, however.

Variance

One limitation of the Markov Chain model is that it doesn't estimate variance. In this section I demonstrate a method for doing so.

I estimated variance by re-sampling the data with replacement (bootstrapping). I took 1000 20 % random samples of households, generating new P^n and X_n matrices (n=1 to 4) for each iteration. Exhibit 12 depicts kernel density plots of the distributions for state1 (normal burden). The predicted share of households with normal burden decreases and estimated variance increases with time.

Exhibit 13-16 reports predictions for 2008 admissions for states 1-4, respectively, along with 95 % confidence intervals. X_0={.613, .331, .028, .030}. The model predicts a pattern of monotonically increasing burden. The share of households with normal burden decreases each year, while the share with medium, high, and extreme burden continually increase. Rates of increase decrease in 2011, however, when burdens start leveling off. Variance for normal burden is smaller than for the other categories, due to the larger number of households with normal burden.

Exhibit 12: Kernel Density Plot of Bootstrap Estimates, State 1 (Normal Burden) Years 2-5

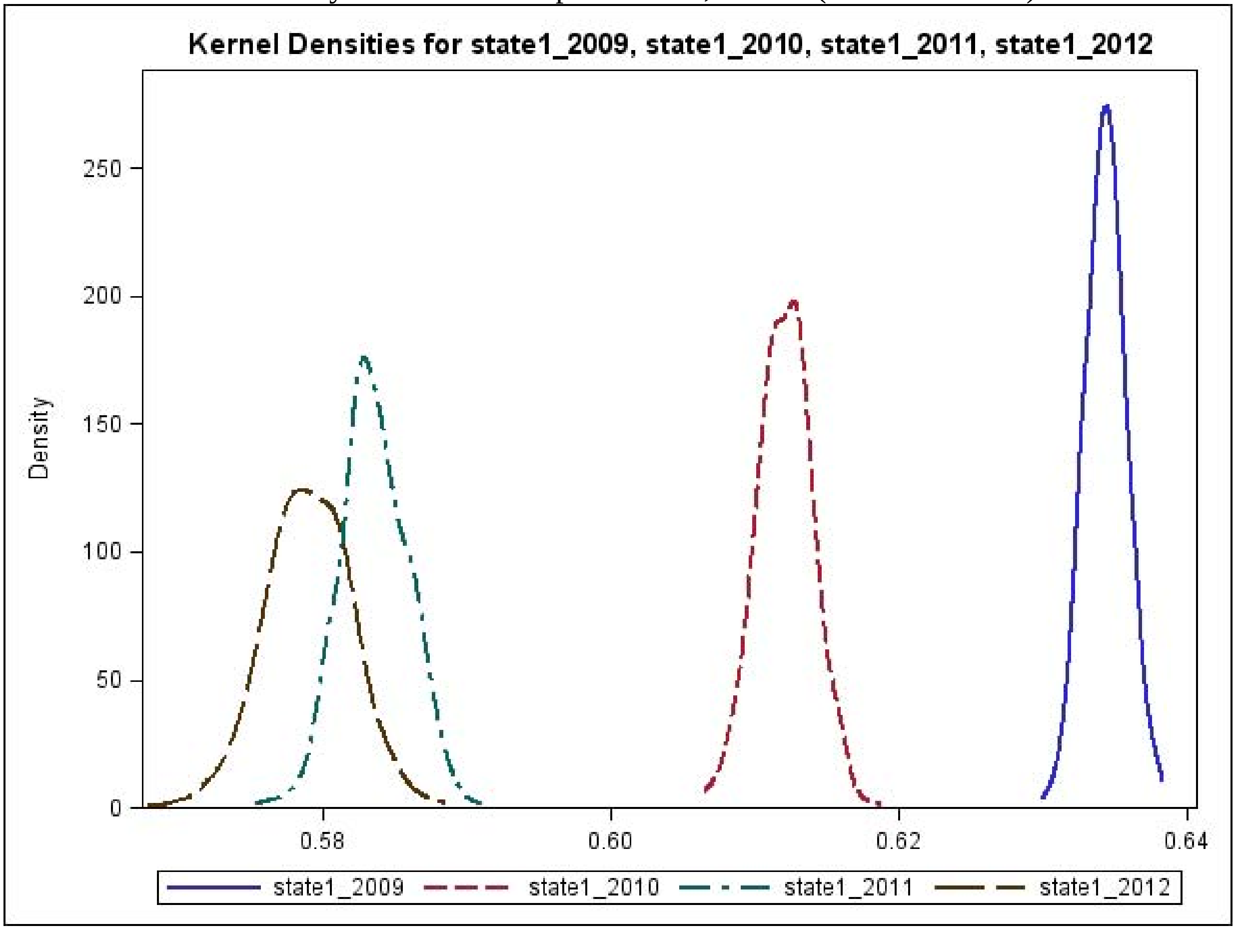


Exhibit 13: Bootstrap Predictions for 2008 Admissions, State 1(Normal Burden)

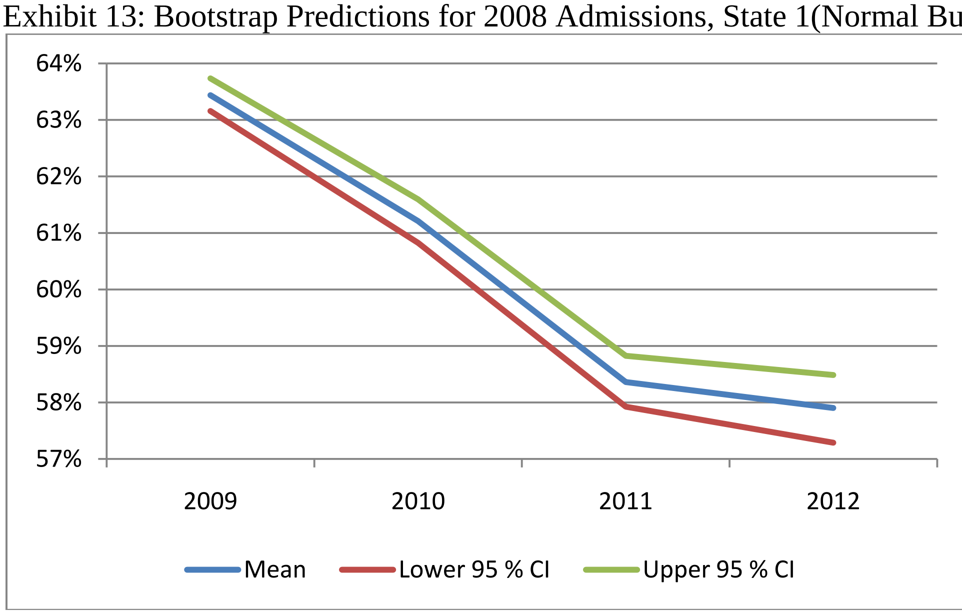


Exhibit 14: Bootstrap Predictions for 2008 Admissions, State 2(Medium Burden)

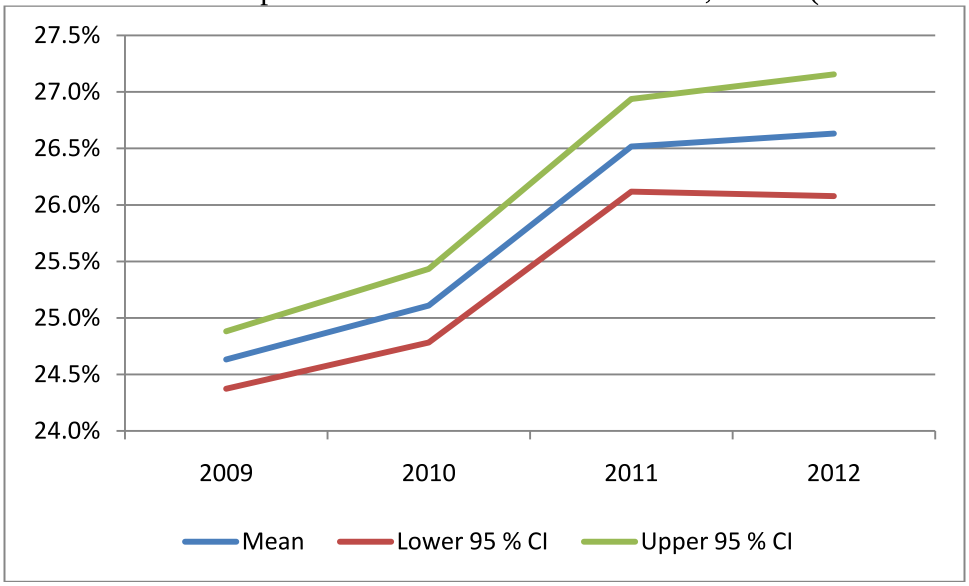


Exhibit 15: Bootstrap Predictions for 2008 Admissions, State 3(High Burden)

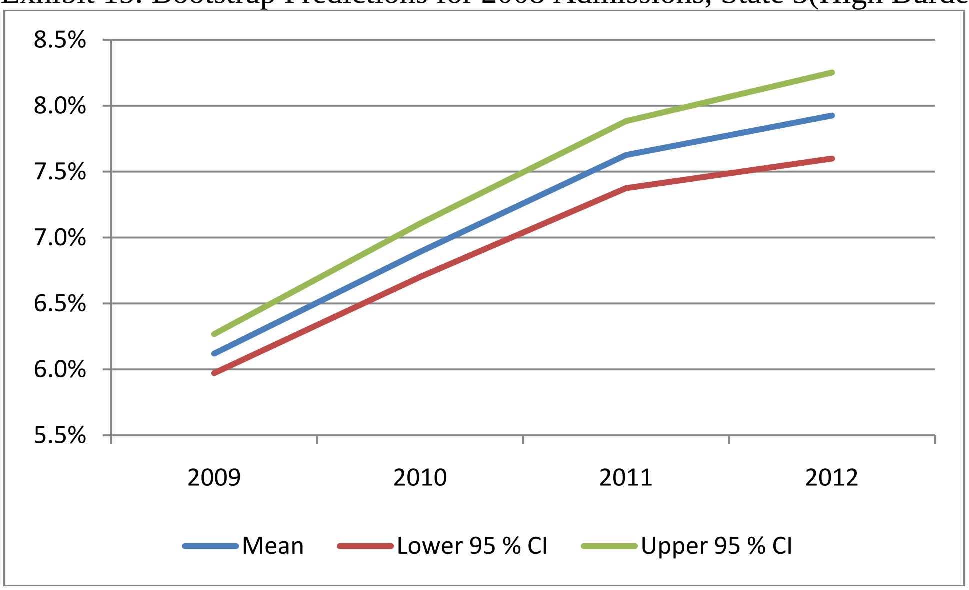
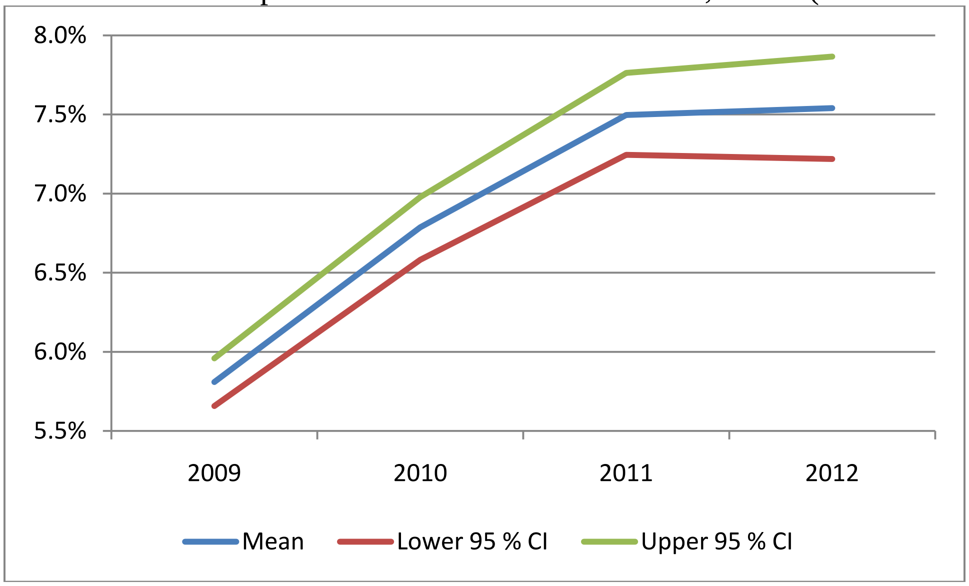


Exhibit 16: Bootstrap Predictions for 2008 Admissions, State 4(Extreme Burden)



VII Conclusion

This study models rent burden in the Housing Choice Voucher Program as a Markov Chain. The model predicts rent burden with program tenure, using longitudinal household data for 2000 through 2008.

Results indicate rent burden increases the first several years after admission. This is perhaps not surprising, given that initial restrictions on burden at admission don't apply in following years.

The model also predicts considerable mobility across households over time. For a family with burden of 50 % or more at admission, their estimated chance of burden below 32 % in their next year is 54 %.

One limitation of the Markov Chain model is that it doesn't estimate variance. I demonstrate a simple method for doing so via bootstrapping.

References

Clark, W. A. V. 1965(June) "Markov Chain Analysis in Geography: An Application to the Movement of Rental Housing Areas" *Annals of the Association of American Geographers*, Vol. 55, No. 2, pp. 351-359. http://www.jstor.org/pss/2561762

Cortes, Alvaro, Ken Lam, and David Fein 2008 "Household Lifecycle and Length of Stay in Housing Assistance Programs" *Cityscape: A Journal of Policy Development and Research* Vol. 10(1). http://www.huduser.org/periodicals/cityscpe/vol10num1/ch5.html

McClure, Kirk 2005 "Rent Burden in the Housing Choice Voucher Program," *Cityscape: A Journal of Policy Development and Research* Vol. 8(2). http://www.huduser.org/Periodicals/cityscpe/vol8num2/ch1.html

Ross, Sheldon M. 2007 *Introduction to Probability Models*, 9th Ed., Academic Press: London.

Susin, Scott 2007 "Duration of Rent Burden as a Measure of Need," *Cityscape: A Journal of Policy Development and Research* Vol. 9(1). http://www.huduser.org/Periodicals/cityscpe/vol9num1/ch7.html

Martin, Marge, Scott Susin, and Barry Steffen. 2005. "Rent Burden: Robustness of Estimates and Duration." In *Affordable Housing Needs: A Report to Congress on the Significant Need for Housing,* edited by the U.S. Department of Housing and Urban Development, Office of Policy Development and Research. Washington, DC. http://www.huduser.org/Publications/pdf/AffHsgNeedsRpt2003.pdf