

Jeremy Yu <jky@u.washington.edu>
Professor Startz
Economics 482 A
Term Paper – Final Draft
2003 December 9

Racial Discrimination in the Provision of Homeownership Opportunities in Seattle

Executive Summary

Widespread racial segregation and discrimination, both past and current, in housing in the United States have been reported in the literature. The disparate homeownership rates of non-Hispanic whites, and blacks and Hispanics, are particularly remarkable. There is a question, however, as to the extent on-going discrimination exists in particular localities. This paper reports on the results of a multiple regression analysis designed to indicate whether racial discrimination exists in the distribution of housing tenure in the city of Seattle. Census block group-level data from the 2000 Census were used. A regression model with the tenure rate of a block group as a function of householder age (the proportion of householders under 35), median household income, the rate of marriage, the proportion of households with children under 18, the proportion of households receiving retirement income, the proportion of householders who are black, and the proportion of householders who are Hispanic, was compared to a restricted regression model of block group homeownership rates that excluded the racial predictors. An unexpected small positive coefficient, not significantly different from zero, on the black householders variable was calculated, but there was a large negative coefficient on the Hispanic householders variable, and the hypothesis that the two coefficients on the racial variables were jointly equal to zero was strongly rejected at the 95% confidence level. The restricted regression

model was found to be significantly different from the unrestricted model. Problems having to do with the interpretation of “tenure” and the “Owner occupied” category in the Census data are discussed.

Introduction

Nation-wide, 72% of non-Hispanic white householders own, rather than rent, their homes while 46% of black householders and 46% of Hispanic householders own their homes (U.S. Census Bureau, 2000a). In Seattle, 52% of non-Hispanic white householders own their homes while only 36% of black householders and 25% of Hispanic householders own their homes.

Racial housing segregation exists and, as racial residential segregation, is also reflected in persistently high to “moderate” dissimilarity indexes nation-wide (National Research Council, 1999) and locally. The black-white dissimilarity index for Seattle in 1990 is 56. Segregation in homeownership status is one expression of housing segregation.

While the existence of de facto racial segregation in housing is not questioned, “many people assert that discrimination is no longer a problem” (Yinger, 2001, p. 372), in other words, that there is currently segregation, but not discrimination. Segregation is assumed to be a consequence of past discrimination alone. Yinger (2001) provides both qualitative and quantitative evidence supporting the existence in the United States of on-going racial discrimination in the housing market, which strongly influences where households are located, homeownership status, and housing segregation in general. In this paper, I look at whether such discrimination exists in the local area.

All else, including poverty, equal, does minority race restrict access to housing opportunities in Seattle? I emphasize access to housing *opportunities*, rather than residential location or another more neutral term, because the existence of de facto segregation is not being

debated. Also, I am not concerned here with the assumed possibility of minority self-segregation. The issue is whether racial minorities are being prevented from having the housing opportunities that they would probably like to have and which they are economically or financially capable of having in terms of their budget. This has historically been the case. For example, in 1970, “Black persons must have a higher income than whites in Seattle to live in the same place” (Seattle Human Rights Department, 1976, p. 13). This was due to a variety of things, such as racial redlining, a practice that continues to this day (Johnson-Elie, 2003).

One desirable housing opportunity is homeownership, which is a form of housing wealth and, importantly, a way of avoiding poverty (Yinger, 2001). And homeownership allows access to suburban jobs and other employment opportunities. Homeownership may also allow access to better schools, or educational opportunities.

I used block group-level 2000 Decennial Census data obtained from the U.S. Census Bureau (U.S. Census Bureau, 2000b) and restricted and unrestricted multiple linear regression models to investigate whether householder race (and Hispanic origin) made a statistically significant contribution to the determination of homeownership by householders, or whether the selected householder age, family-related and household income indicators by themselves were adequate to predict homeownership. A Wald test was done to determine whether the restricted and unrestricted models were significantly different.

Data

My decision to use Census data was based on the easy accessibility of the data, the large size of the available data, and the consistency with which the data were generated. The decision to use (census) block group-level data was based on the need to use the lowest level of aggregation possible without raising the issue of suppressed data. Also, income indicator data

were not at all supplied for individual blocks. (Summary File 3, based on the Census long-form sample, does not include block-level data. Block-level data would have highly unreliable confidence intervals as well as suppressed data.) Additionally, block group-level data, which represented only a few street block-size geographies, represented households in the same neighborhoods, which was useful for my particular analysis. It was easier for the purpose of interpretation to think about the homeownership rates of clusters of households, rather than of formally constructed groups of spatially dispersed households (less meaningful).

In the 2000 Census, there are a total of 570 block groups for the city of Seattle. Not interested in differences between districts or between any other geographies smaller than the city, I included all of the block groups for Seattle in my analysis. One block group, for which there were no data (no households), was excluded ($n = 570 - 1 = 569$).

I hypothesized that race aside, homeownership, as the proportion of occupied housing units in a block group that were owner-occupied, was determined mainly by the proportion of householders younger than 35 years old (in the block group), median household income, the proportion of households that were married-couple family households, the proportion of households in whom own children younger than 18 years old were present, and the proportion of households receiving retirement income. I then considered race, and Hispanic origin, and decided to include in the unrestricted regression model the proportion of householders who were black (non-Hispanic black, to decrease correlation with HISPANIC) and the proportion of householders who were Hispanic. The proportion of householders who were non-Hispanic white was excluded to limit multicollinearity. Proportions and other relative figures (such as median, not aggregate, household incomes), rather than counts, or absolute numbers, were included wherever possible in the regression models and represented my assumption that the independent

variables determined the proportion, rather than the absolute number, of occupied housing units in a block group that were owned. As block groups varied in the number of households that they contained, this assumption was appropriate.

The selected variables were coded as follows in the Census 2000 Summary File 3 (SF 3).

- P10. HOUSEHOLD SIZE BY HOUSEHOLD TYPE BY PRESENCE OF OWN CHILDREN UNDER 18 YEARS [19] - Universe: Households
- P13. HOUSEHOLD TYPE BY AGE OF HOUSEHOLDER [19] - Universe: Households
- P53. MEDIAN HOUSEHOLD INCOME IN 1999 (DOLLARS) [1] - Universe: Households
- P65. RETIREMENT INCOME IN 1999 FOR HOUSEHOLDS [3] - Universe: Households
- H7. TENURE [3] - Universe: Occupied housing units
- H10. HISPANIC OR LATINO HOUSEHOLDER BY RACE OF HOUSEHOLDER [17] - Universe: Occupied housing units

The numbers between brackets were the numbers of variables, or fields, within each matrix, or table. For example, matrix H7 contained the three fields named “Total,” “Owner occupied,” and “Renter occupied.” In my analysis, the homeownership rate of a geography, here a block group, was represented by the second field, “Owner occupied,” divided by the first, “Total.” Other variables included in the regression models were similarly constructed. The unrestricted regression model was expressed in terms of the SF 3 tables and then their fields, in brackets. The first field of a table was indexed by the number 1.

$$\begin{aligned}
 H7[2] \div H7[1] = & \beta_0 + \\
 & \beta_1(P10[7] \div P10[1]) + \\
 & \beta_2((P10[8] + P10[12] + P10[15]) \div P10[1]) + \\
 & \beta_3((P13[3] + P13[4] + P13[12] + P13[13]) \div P13[1]) + \\
 & \beta_4(P53[1]) + \\
 & \beta_5(P65[2] \div P65[1]) + \\
 & \beta_6((H10[4] + H10[12]) \div H10[1]) + \\
 & \beta_7(H10[10] \div H10[1]) + \\
 & \varepsilon
 \end{aligned}$$

The model was more simply represented as

$$\text{TENURE} = \beta_0 + \beta_1\text{MARRIED} + \beta_2\text{CHILDREN} + \beta_3\text{AGE35} + \beta_4\text{INCOMEM} + \beta_5\text{RETIRED} + \beta_6\text{BLACKNH} + \beta_7\text{HISPANIC} + \varepsilon.$$

The restricted regression model, without the racial variables, was represented as

$$\text{TENURE} = \beta_0 + \beta_1\text{MARRIED} + \beta_2\text{CHILDREN} + \beta_3\text{AGE35} + \beta_4\text{INCOMEM} + \beta_5\text{RETIRED} + \varepsilon.$$

Results

Tabulated results of the regression of TENURE on the independent variables of the unrestricted regression model are presented in Appendix A. The results for the restricted regression model are presented in Appendix B.

In a preliminary unrestricted regression that did not specify RETIRED as a regressor, a negative coefficient (-0.016 and not significantly different from zero) on CHILDREN was calculated. The negative sign of the coefficient was unexpected but might have been explained by the existence of non-married family households with female householders, and the proportion of these with own children younger than 18 years old. Compared to other types of households, I did expect these households to be located in owned (owner-occupied) housing units. After RETIRED was included as a regressor, the coefficient on CHILDREN became positive (0.052 but still not significantly different from zero), and the coefficient (0.31 and significantly different from zero) on RETIRED was positive. The sign change expresses that CHILDREN reflects RETIRED, in other words, that the proportion of households receiving retirement income is

inversely related to the proportion of households with own children younger than 18. However, all else equal, TENURE increases with CHILDREN.

The negative coefficient (-0.38 in the unrestricted model) on AGE35 was not surprising. The greater the proportion of households with householders younger than 35, the lower the tenure rate. A positive coefficient could have been obtained by setting $AGE35 = 1 - AGE35$.

The negative and comparatively large coefficient (-0.44) on HISPANIC strongly agreed with the hypothesis of racial discrimination in the provision of homeownership opportunities. However, there was a small but unexpected positive coefficient (0.065) on BLACKNH. It appeared that all else equal, tenure rate was positively correlated with the proportion of households with black householders. However, the coefficient on BLACKNH was not significantly different from zero, and any number of negative values, at the 95% confidence level.

As a whole, the selected racial variables contributed significantly to the prediction of block group homeownership rate. At the 95% confidence level, the hypothesis that the coefficients on BLACKNH and HISPANIC were jointly zero was strongly rejected ($P = 0.0067$). (Computer-generated output of the Wald test for the joint equality is presented in Appendix C.) The unrestricted regression model was significantly different from the restricted model.

Discussion

To the extent that determination by racial variables after adjustment for other, non-racial factors represents racial discrimination, the results support the conclusion that racial discrimination exists in the provision of homeownership opportunities in Seattle. Not only did the unrestricted regression model provide a significantly better fit to the observed values of model variables, the presence of racial minority householders was, all else equal, negatively correlated with tenure rate in a block group. However, while small and statistically insignificant,

the positive coefficient on the black householders variable is not desirable from the point of view of consistency and needs to be investigated. I speculate that the positive correlation is a consequence of the quality composition of the actual owner-occupied housing units in Seattle, the pressuring of racial minorities into lower-quality owned units, and the restriction of higher-quality rented units to whites. In fact, the median gross rents, which represent rented housing unit values, for blacks and Hispanics in Seattle, \$616 and \$671, are significantly less than the median gross rent for whites in Seattle, \$752 (U.S. Census Bureau, 2000b).

It can be reasonably assumed that owned units and rented units are each relatively homogenous, and that owning is generally preferable to renting, but the “owner occupied” and “renter occupied” categories of the “Tenure” housing table in SF 3 obscure the possibility that some renter-occupied units are preferable to some owner-occupied units.

References

- Johnson-Elie, T. (2003, October 28). Home loan disparities between blacks, whites tell familiar tale. *Milwaukee Journal Sentinel*. Retrieved November 8, 2003, from <http://www.jsonline.com/bym/Biz2biz/oct03/180824.asp>
- National Research Council Committee on Improving the Future of U.S. Cities through Improved Metropolitan Area Governance. (1999). *Governance and opportunity in metropolitan America* [Electronic version]. Washington, D.C.: National Academy Press.
- Seattle Human Rights Department. (1976). A study and data on segregated housing in Seattle, Washington (H-3721). City of Seattle: Seattle.
- U.S. Census Bureau. (2000a). *Census 2000 Summary File 2*. Retrieved November 8, 2003, from the American FactFinder:
http://factfinder.census.gov/servlet/DatasetMainPageServlet?_ds_name=DEC_2000_SF3_U&_program=DEC&_lang=en
- U.S. Census Bureau. (2000b). *Census 2000 Summary File 3*. Retrieved November 8, 2003, from the American FactFinder:
http://factfinder.census.gov/servlet/DatasetMainPageServlet?_ds_name=DEC_2000_SF3_U&_program=DEC&_lang=en
- Yinger, J. (2001). Housing discrimination and residential segregation as causes of poverty. In S. H. Danziger & R. H. Haveman (Eds.), *Understanding Poverty* (pp. 359-391). New York: Russell Sage Foundation.

Appendix A

Unrestricted regression

Dependent Variable: TENURE
 Method: Least Squares
 Date: 11/24/03 Time: 10:33
 Sample: 1 570
 Included observations: 569
 Excluded observations: 1

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.182779	0.036597	4.994341	0.0000
MARRIED	0.651229	0.067647	9.626919	0.0000
CHILDREN	0.051867	0.071786	0.722519	0.4703
AGE35	-0.380538	0.052373	-7.265880	0.0000
INCOMEM	3.55E-06	3.43E-07	10.36032	0.0000
RETIRED	0.309103	0.090868	3.401665	0.0007
BLACKNH	0.065406	0.051247	1.276281	0.2024
HISPANIC	-0.435702	0.148871	-2.926712	0.0036
R-squared	0.805835	Mean dependent var	0.547628	
Adjusted R-squared	0.803412	S.D. dependent var	0.267370	
S.E. of regression	0.118547	Akaike info criterion	-1.413049	
Sum squared resid	7.883993	Schwarz criterion	-1.351976	
Log likelihood	410.0126	F-statistic	332.6127	
Durbin-Watson stat	1.479755	Prob(F-statistic)	0.000000	

Estimation Equation:

=====

$$\text{TENURE} = C(1) + C(2)*\text{MARRIED} + C(3)*\text{CHILDREN} + C(4)*\text{AGE35} + C(5)*\text{INCOMEM} + C(6)*\text{RETIRED} + C(7)*\text{BLACKNH} + C(8)*\text{HISPANIC}$$

Substituted Coefficients:

=====

$$\text{TENURE} = 0.1827792274 + 0.6512286059*\text{MARRIED} + 0.05186698793*\text{CHILDREN} - 0.3805377548*\text{AGE35} + 3.552236685\text{e-}06*\text{INCOMEM} + 0.3091025646*\text{RETIRED} + 0.06540595589*\text{BLACKNH} - 0.4357018362*\text{HISPANIC}$$

The C(1) coefficient, the constant 0.18, suggests an 18% base tenure rate regardless of marriage, children, etc. The C(5) coefficient, on INCOMEM, represents that for every \$10,000 increase in median household income, there is an increase of 0.036 (3.6%) in the tenure rate all else equal. The other coefficients may be interpreted as follows. For every increase of 0.01 (or 1%) in the marriage rate, there is an increase of 0.0065 (0.65%) in the tenure rate, all else equal. So, a block group with a 100% marriage rate would have a much greater tenure rate than an

otherwise similar block group with a 0% marriage rate, by 0.65 (65%). Of course, it is not possible to have a tenure rate greater than 1 (100%), or a tenure rate less than 0 for that matter. So a block group with a 0% marriage rate but a tenure rate of, say, 75%, would not, in reality, have a tenure rate of 140% (75% + 65%) if its marriage rate were 100% (but this marriage rate combined with other factors would likely bring the tenure rate close to 100%). Overall, however, the relationship holds, and it is interesting to note that an all-Hispanic (HISPANIC = 1, or 100%) block group would have a much lower tenure rate than an otherwise similar block group, by 0.43 (43%).

Appendix B

Restricted regression

Dependent Variable: TENURE
 Method: Least Squares
 Date: 11/24/03 Time: 11:50
 Sample: 1 570
 Included observations: 569
 Excluded observations: 1

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.172024	0.034224	5.026356	0.0000
MARRIED	0.635006	0.063367	10.02109	0.0000
CHILDREN	0.046784	0.062248	0.751585	0.4526
AGE35	-0.400056	0.051718	-7.735318	0.0000
INCOMEM	3.75E-06	3.29E-07	11.39384	0.0000
RETIRED	0.327404	0.091223	3.589071	0.0004
R-squared	0.802334	Mean dependent var		0.547628
Adjusted R-squared	0.800579	S.D. dependent var		0.267370
S.E. of regression	0.119398	Akaike info criterion		-1.402212
Sum squared resid	8.026124	Schwarz criterion		-1.356407
Log likelihood	404.9293	F-statistic		457.0484
Durbin-Watson stat	1.458884	Prob(F-statistic)		0.000000

Estimation Equation:

=====

$$\text{TENURE} = C(1) + C(2)*\text{MARRIED} + C(3)*\text{CHILDREN} + C(4)*\text{AGE35} + C(5)*\text{INCOMEM} + C(6)*\text{RETIRED}$$

Substituted Coefficients:

=====

$$\text{TENURE} = 0.1720239897 + 0.6350056556*\text{MARRIED} + 0.04678449868*\text{CHILDREN} - 0.4000556187*\text{AGE35} + 3.751571393\text{e-}06*\text{INCOMEM} + 0.327404101*\text{RETIRED}$$

Appendix C

Wald Test:
Equation: Untitled

Test Statistic	Value	df	Probability
F-statistic	5.056795	(2, 561)	0.0067
Chi-square	10.11359	2	0.0064

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(7)	0.065406	0.051247
C(8)	-0.435702	0.148871

Restrictions are linear in coefficients.

Appendix D

Correlation Matrix

	AGE35	BLACKNH	CHILDREN	HISPANIC	INCOMEM	MARRIED	RESID	RETIRED	TENURE
AGE35	1	-0.075986	-0.536393	0.191896	-0.496941	-0.708723	-3.04E-15	-0.620476	-0.737677
BLACKNH	-0.075986	1	0.284116	0.226816	-0.290377	-0.146554	-1.84E-15	-0.046345	-0.111046
CHILDREN	-0.536393	0.284116	1	0.081664	0.317308	0.671613	-3.56E-15	0.237912	0.534785
HISPANIC	0.191896	0.226816	0.081664	1	-0.397418	-0.23674	-3.12E-15	-0.211717	-0.328007
INCOMEM	-0.496941	-0.290377	0.317308	-0.397418	1	0.689601	-5.55E-15	0.352228	0.740707
MARRIED	-0.708723	-0.146554	0.671613	-0.23674	0.689601	1	-6.86E-15	0.548664	0.843634
RESID	-3.04E-15	-1.84E-15	-3.56E-15	-3.12E-15	-5.55E-15	-6.86E-15	1	-2.71E-15	0.440642
RETIRED	-0.620476	-0.046345	0.237912	-0.211717	0.352228	0.548664	-2.71E-15	1	0.569741
TENURE	-0.737677	-0.111046	0.534785	-0.328007	0.740707	0.843634	0.440642	0.569741	1

RESID, the residuals, were generated from the unrestricted regression. The positive correlation of RESID with TENURE indicates that as observed tenure rate increases, predicted tenure rate tends to be further off in the positive direction (more erroneous). As predicted tenure rate increases, observed tenure rate, or actual tenure rate if the regression model is to be used for forecasting, tends to be further off in the negative direction.