Labor and Housing Market Dynamics

Samuel Yandrofski[1]

Abstract

A classic theorem of labor economics is that increased worker mobility causes better employment outcomes. We test this prediction by studying the effect of home ownership on unemployment. Three models are employed in this exercise, with the third producing more reliable estimates than the first two. We attempt to place bounds on the estimates from the third model, and conclude that non-homeowners enjoy superior performance in the labor market vis-à-vis homeowners.

[1] The author thanks Josh Kinsler, David Slichter, Sulagna Mookerjee and Andrew Davis for being generous with their time and providing wonderful comments.

1. Introduction

Governments throughout the world provide favorable tax treatment to individuals who choose to buy homes. One argument in support of this policy is that home ownership produces positive externalities, which causes the amount of home ownership to be below the societally optimal quantity. We do not attempt to document the positive and negative externalities related to home ownership in this paper; instead, we focus on the effect home ownership has on employment outcomes. Understanding the answer to this question is crucial for designing effective tax policies: if home ownership causes poor performance in the labor market, then subsidizing it may be inappropriate.

Coulson and Fisher (2002) and Munch, Rosholm and Svarer (2007) consider similar research questions, and both conclude that home ownership leads to improved performance in the labor market. This paper builds on the work done by these authors by using a different econometric approach to estimate the effect of home ownership on unemployment. A critical assessment of the reliability of various models is also included.

We build three econometric models using micro-level data on individuals from 1985-1995. A multiple regression model with interaction terms is presented first, and is followed by a regression using panel data. The last model uses an instrumental variable to eliminate any bias introduced by reverse causality. The strengths and weaknesses of each model are discussed, and the paper concludes with comments on the implications for public policy.

2. Summary Statistics

Before discussing the details of the models, it is helpful to get a sense of the data by considering the following table of summary statistics.

	All individuals		**Non-homeowners**		**Homeowners**	
Variable	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Number of weeks unemployed	2.52	7.40	3.28	8.40	1.21	4.98
Age	28.42	3.51	27.68	3.46	29.72	3.22
Cognitive ability	43.22	28.34	38.43	27.71	51.52	27.50
HS dropout	0.11	0.31	0.13	0.34	0.07	0.25
HS graduate	0.44	0.50	0.44	0.50	0.45	0.50
Some college	0.24	0.43	0.24	0.43	0.25	0.43
College graduate	0.21	0.41	0.19	0.39	0.24	0.43
Female	0.51	0.50	0.49	0.50	0.55	0.50
Hispanic	0.16	0.37	0.18	0.38	0.14	0.35
Black	0.27	0.44	0.34	0.47	0.15	0.35
Married	0.51	0.50	0.33	0.47	0.82	0.38
Children	0.51	0.50	0.40	0.49	0.69	0.46
Low home ownership as child	0.24	0.43	0.26	0.44	0.19	0.40
Med home ownership as child	0.52	0.50	0.51	0.50	0.52	0.50
High home ownership as child	0.25	0.43	0.22	0.42	0.29	0.45

Figure 1.

From these summary statistics we see that homeowners and non-homeowners differ in important ways. During the years 1985-1995, the average non-homeowner spent 3.28 weeks unemployed per year, while the average homeowner spent only 1.21 weeks unemployed per year. Homeowners also tended to be older, more educated and substantially more likely to be married than their non-homeowner counterparts.

3.1. Multiple Regression

We begin by estimating a multiple regression model with several interaction terms. The model takes the following form; a table with a description of each regressor is included on page 13.

$$weeks\_unemployed_i = \beta_0 + \beta_1 own_i + \beta_2 age_i + \beta_3 TestPCT_i + \beta_4 HsDrop_i + \beta_5 SC_i + \beta_6 Coll_i + \beta_7 female_i + \beta_8 hispanic_i + \beta_9 black_i + \beta_{10} married_i + \beta_{11} children_i + \beta_{12} ownAge_i + \beta_{13} ownTestPCT_i + \beta_{14} ownHSDrop_i + \beta_{15} ownSC_i + \beta_{16} ownColl_i + u_i$$

Each regressor is included to reduce the bias in the estimates of the coefficient on *own*. The nonlinear terms are introduced to capture any interactions between the relevant regressors: perhaps home ownership does reduce worker mobility, but reduced mobility is particularly deleterious for workers with little human capital.

The multiple regression model has serious limitations, and two of them are briefly described here. First, there are variables missing from the regression that might cause omitted variable bias. The regression does not distinguish between workers living in a metropolitan statistical area (MSA) and those living in less urban settings. It is obvious that living in an MSA and home ownership are negatively correlated, and it is plausible that the diversity of industries found in many MSAs could help workers adjust to adverse demand shocks.

The results from the multiple regression model indicate that there is a negative and statistically significant relationship between unemployment and home ownership. (A complete summary of the results is available on page 6.) But the model does little to

show which direction causation runs: does home ownership reduce unemployment, or does low unemployment cause home ownership?

3.2. Fixed Effects

A fixed effects regression model is also employed in the quest to estimate the effect of home ownership on unemployment. In this model, $Weeks\_unemployed_{it}$ represents the number of weeks worker *i* was unemployed in year *t*. The regressors are defined similarly.

$$Weeks\_unemployed_{it} = \beta_0 + \beta_1 own_{it} + \beta_2 age_{it} + \beta_3 married_{it} + \beta_4 children_{it} + \beta_5 ownAge_{it}$$

$$+ \beta_6 ownTestPct_{it} + \beta_7 ownHsDrop_{it} + \beta_8 ownSC_{it} + \beta_9 ownColl_{it} + \alpha_i + u_{it}$$

One might argue that unobservable characteristics of workers – such as preferences for leisure or levels of motivation – are time-invariant, correlated with home ownership and significant causes of labor market performance. If this is true, the fixed effects regression model will eliminate any omitted variable bias coming from these excluded variables. However, it is also plausible that individuals become more motivated and socially adept as they age. These variables are important determinants of unemployment outcomes; if they are time-variant, fixed effects regression cannot control for them and the estimates of the coefficient on home ownership will be biased.

The data set used to generate this model does not contain information on whether workers reside in urban areas, and the exclusion of a variable measuring the effect of living in an MSA is another potential for omitted variable bias.

The results from the fixed effects model are included on page 7. The relationship between home ownership and unemployment continues to be negative and statistically

significant, but the magnitude of the coefficient on home ownership has increased from the multiple regression model.

In general, the fixed effects regression model has difficulty differentiating causation from correlation. This concern is particularly pronounced in the context of the dynamic between employment outcomes and home ownership: it is difficult to imagine that the decision to purchase a home is not tied in some way to labor market performance. This endogeneity problem makes it impossible to claim causality or propose policy recommendations.

Dependent variable: weeks of unemployment (Multiple Regression)					
	[1]	[2]	[3]	[4]	[5]
Intercept	**3.27**	**4.363**	**4.399**	**4.57**	**4.708**
Home owner	**-2.07**	**-1.644**	**-1.667**	**-1.01**	-1.007*
Age		0.008	0.006	-0.001	0.006
Score on cognitive test		**-0.034**	**-0.028**	**-0.019**	**-0.026**
High school dropout			0.229	**0.356**	**0.398**
Some college			**-0.542**	**-0.666**	**-0.743**
College graduate			**-0.56**	**-0.838**	**-0.824**
Female				**-0.549**	**-0.543**
Hispanic				**-0.241**	**-0.238**
Black				**1.108**	**1.052**
Married				**-1.054**	**-1.051**
Children				0.05	0.029
Home owner * Age					-0.031
Home owner * Test score					**0.019**
Home owner * HS dropout					-0.485*
Home owner * Some college					0.241
Home owner * College					0.003
R^2	.0181	.0339	.0352	.0456	.0471

Bold denotes significant at the 5% level

* denotes significant at the 10% level

Figure 3.

Dependent variable: weeks of unemployment (Fixed Effects)				
	[1]	[2]	[3]	[4]
Intercept	**2.745**	**4.706**	**4.610**	**5.019**
Home owner	**-0.614**	**-0.41**	**-0.309**	**-1.812**
Age		**-0.072**	**-0.06**	**-0.0750**
Married			-0.229*	-0.216*
Children			**-0.315**	**-0.338**
Home owner * Age				0.043*
Home owner * Test Score				0.007
Home owner * High school dropout				-0.531
Home owner * Some college				0.025
Home owner * College				-0.255

Bold denotes significant at the 5% level
* denotes significant at the 10% level

Figure 4.

3.3. Instrumental Variables Regression

The possibility of simultaneous causality along with lingering concerns about omitted variables bias leads us to consider using a third type of econometric model. We consider a regression estimated by the method of instrumental variables; the home ownership rate in an individuals' home county is the proposed instrument. We begin by presenting the model, and then discuss whether conditions of instrument validity have been satisfied.

Two-stage least squares (TSLS) is used as an estimator, and the reduced form equation for *own* becomes

$$own_i = \gamma_0 + \gamma_1 med\_own_i + \gamma_2 high\_own_i + \gamma_3 female_i + \gamma_4 married_i + v_i. \quad (1)$$

The variables *female* and *married* are assumed to be exogenous (an assumption that will be examined more closely later). The regressor *low_own* is excluded to avoid perfect multicollinearity.

The second stage of TSLS estimates

$$weeks\_unemployed_i = \beta_0 + \beta_1 \hat{own}_i + \beta_2 female_i + \beta_3 married_i + u_i, \quad (2)$$

where $\hat{own}_i$ is an estimate obtained from the first stage of TSLS.

Variables measuring education, cognitive ability and race are omitted from both (1) and (2) because they are correlated with unobserved factors that affect labor market performance. For example, education is correlated with family structure, and it is likely that family structure has an impact on employment outcomes. The estimation results from TSLS are included below.

Dependent variable: weeks of unemployment (IV)	
	[1]
Intercept	**3.292**
Home owner	**3.873**
Female	**-0.647**
Married	**-3.630**
Instrumental Variable	ownership rate in home county
First-stage F-statistic	131.72

Bold denotes significant at the 5% level
* denotes significant at the 10% level

Figure 5.

The results from the instrumental variables regression differ substantially from those generated by multiple regression and fixed effects: the sign of the coefficient on *own* has gone from negative to positive and the magnitude has increased. However, we have not yet verified whether the instrument is valid, and an acceptance of the results should be postponed until a more thorough analysis has been performed. Two conditions are needed for an instrument to be valid: relevance and exogeneity. We look at each of these conditions in turn.

The first stage F-statistic tests for correlation between the respondents' decision to buy a home and the prevalence of home ownership in the respondents' home county. The F-statistic is 131.72, which is well above the commonly accepted cutoff of 10. This provides strong evidence that the condition of instrumental relevance has been satisfied.

There is less evidence that the condition of exogeneity has been satisfied. A respondent who was raised in a county with a high percentage of home ownership is likely to be different from a respondent raised in a county with a low percentage of home ownership in important ways. For example, individuals born in a county with a high

percentage of home ownership probably also have parents who have been successful in the labor market. Because the success of parents in the labor market is likely to be correlated with the success of their children in the labor market, the exogeneity condition fails.

Although the failure of the exogeneity condition causes the instrument to be invalid, this does not mean the instrumental variables technique cannot produce any insights about the research question. In the subsequent paragraphs, we produce an argument that attempts to place a lower bound on the coefficient on home ownership.

There are several variables in the error term; we list the top seven here in order of importance.

(1) Motivation
(2) Quality of post-secondary education
(3) Quantity of education
(4) Social skills
(5) Unemployment rate in home state
(6) Family structure
(7) Success of parents in the labor market

Let Z denote percentage of home ownership in home county, y denote weeks of unemployment, x denote home ownership, u_1 denote motivation, u_2 denote quality of post-secondary education, and so on.

cov(Z, u_1) = indeterminate	cov(u_1, y) = negative
cov(Z, u_2) = positive	cov(u_2, y) = negative
cov(Z, u_3) = positive	cov(u_3, y) = negative
cov(Z, u_4) = indeterminate	cov(u_4, y) = negative
cov(Z, u_5) = positive	cov(u_5, y) = negative
cov(Z, u_6) = positive	cov(u_6, y) = negative
cov(Z, u_7) = positive	cov(u_7, y) = negative

Figure 6.

This analysis demonstrates that the instrument affects labor market performance through channels other than home ownership. But note that for each u_i with cov(Z, u_i) > 0, we have cov(u_i, y) > 0. This indicates that our estimate of the coefficient on home ownership is artificially low; we can therefore use the estimate obtained from the instrumental variables regression as a lower bound on the coefficient on home ownership. Because our estimate of the effect of home ownership on unemployment was positive and statistically significant, this provides strong evidence that non-homeowners experience better outcomes in the labor market than their homeowner counterparts.

3. Conclusion

Several models have been employed to estimate the effect home ownership has on labor market performance. Regressions using OLS and panel data produced very different estimates than the method of instrumental variables. The first two models struggled to differentiate causation from correlation, whereas the latter was able to make significant progress in this respect. The stark differences in results indicate that causality runs in both directions between home ownership and unemployment.

Because of its ability to deal with omitted variable bias and simultaneous causality, the method of instrumental variables was the most reliable econometric approach. An instrumental variable with a higher degree of exogeneity would increase our confidence in the predictions made about the dynamic between home ownership and unemployment. In the absence of a perfectly exogenous instrument, we resorted to an analysis of the composition of the error term and introduced a lower bound on the coefficient of interest. Researchers who are able to control for some of the confounding variables that caused our instrument to fail the exogeneity condition would likely find fertile research ground going forward.

Classic models of labor economics posit that increased worker mobility is associated with better employment outcomes. Our results agree with this theory: purchasing a home causes workers to become less mobile, and a price is paid in the labor market for this reduced mobility.

The results presented here do not necessarily imply that subsidies for home ownership should be reduced. An exhaustive catalog of the positive and negative externalities associated with home ownership would need to be compiled before this type of recommendation could be made. However, there is empirical evidence that employment prospects are reduced by home ownership, and policy makers should consider the negative externalities associated with reduced employment prospects when designing future housing policies.

Regressor	Description
own	Binary variable with value 1 if worker owns home
age	Age of worker in years
TestPCT	Score on test that measures cognitive ability
HsDrop	Binary variable with value 1 if worker is a high school dropout
SC	Binary variable with value 1 if worker has completed some college
Coll	Binary variable with value 1 if worker is a college graduate
Female	Binary variable with value 1 if worker is female
Hispanic	Binary variable with value 1 if worker is Hispanic
Black	Binary variable with value 1 if worker is Black
Married	Binary variable with value 1 if worker is married
Children	Binary variable with value 1 if worker has one or more children
ownAge	Interaction term defined as own * age
ownTestPCT	Interaction term defined as own * TestPCT
ownHsDrop	Interaction term defined as own * HsDrop
ownSC	Interaction term defined as own * SC
ownColl	Interaction term defined as own * Coll
low_own	Binary variable with value 1 if home ownership rate in worker's home county is less than 60%
med_own	Binary variable with value 1 if home ownership rate in worker's home county is between 60% and 75%
high_own	Binary variable with value 1 if home ownership rate in worker's home county is greater than 75%

Note: High school graduates, males, whites, singles, workers without children and workers from counties with low rates of home ownership are the excluded categories for the relevant regressors.