

Estimating Interstate Differences in Homeless Population:

Aggregate Study

ECON 4319

Yuya Ogawa

Table of Contents

Introduction	3 - 4
Literature Review	4 - 6
Data Description and Methodology	6 - 10
Conclusion	10 - 11
Appendix	12
Works Cited	13 - 14

Introduction

When you go into the city area, aside from skyscrapers and businesspeople, you would inevitably encounter homeless people, whereas, if you were in the country area, you would hardly see them on the street. In Japan, a country with 126 million people has only 4,000 homeless people according to the Japanese Ministry of Health, Labor, and Welfare, whereas in the United States, a country with 330 million people, there are about 553,000 homeless population according to the Department of Housing and Urban Development. The homeless per capita is over five times greater in United States than in Japan. This is a development question since it relates to key idea of Amartya Sen's capability approach to development in the sense that the question of homelessness concerns with things that are apart from income that decides people's wellbeing such as housing.

Why do some regions have less homeless population than other regions? In this paper, we are interested in finding an answer to this perplexing regional differences. In other words, we are going to look at the macroeconomic factor that can influence the homeless population, rather than microeconomic choices that homeless people make. To do this, we will aggregate macro-level data to explain the state-wide regional differences in homeless population. First, we are interested in finding the homeless per capita ratio for each state in the United States, and run a multilinear regression using saving rate, housing cost as a percentage of income, and inequality in the area. The result of the regression is a linear probability model with the coefficients representing probability.

Our result tells us that the housing price per income and the saving rate of those that live below poverty level is a significant factor in predicting homelessness. Moreover, if we take into account the effect of inequality and saving in combination, we found that all them are significant

factor in predicting interstate homeless ratio. Our result highlights the importance of equality and asset level especially for those that live below poverty level and points us to the direction of addressing extreme inequality issue in the city area.

Literature Review

Study of homelessness can be more or less be divided into two subjects: microeconomic factors of homelessness and macroeconomic factors. Microeconomic factors focus on the individual case studies and evaluate the effect of policies that are designed to help homeless population in the individual levels. However, in macroeconomic level studies, the study of interest includes aggregate factors that drives homelessness such as unemployment rate, wage, and so on. For instance, O’Flaherty has approached the notion of shelter quality and living conditions using the traditional concept of “demand” and “supply.” Precisely, he has introduced a new perspective in which you would regard the demand (or utility) of becoming a homeless determined by the shelter conditions and living conditions of street. Furthermore, analysis becomes rather interesting once incorporating the “generosity” as determined by the perspective or attitude toward homeless people (O’Flaherty 2019). This has a profound implication in the sense that reduction in homeless population can be jointly determined by the “generosity” curve and “demand” curve, and individual case research as well as aggregate research must incorporate this fact while identifying the causality.

Indeed, weather conditions, poverty rates, and religiosity were significant factor in explaining the varying population of homelessness between cities which verifies the O’Flaherty’s claim (Corinth and Lucas 2018). Thus, most of our research in finding the factor of homelessness partially has to do with identifying the variables that impacts the “generosity”

curve and “need” curve (O’Flaherty 2019). However, our research of interest deals with external economic situations that can neither influence the homelessness’s need curve nor the generosity curve. Hence, the study of macroeconomic factors is still important consideration in homelessness research.

To discuss some of the underlying macroeconomic factors determining homelessness, Corinth has looked at the causal impact of permanent supportive housing on homelessness through quasi-experimental method and concluded that housing support indeed reduces homeless population (Corinth 2017). This resolves the inverse causality claim and clears some of the misconceptions regarding homelessness as a determinant of higher housing prices. Thus, our research must incorporate the housing as one of the explanatory variables.

By far we have discussed the necessity of considering microeconomic factor as well as macroeconomic factors such as housing price, but we would like to now discuss some of the literature that looked at the interstate differences in homeless population. For instance, Cebula and Alexander have attempted to answer the interstate difference by performing a multilinear regression on the homeless count including the following variables: cost of living index in state, percentage of the population in state age 25 years and above without a high school diploma, percentage of the population in state age 25 years and above with a bachelor’s degree or higher, index of labor market freedom, the level of per capita personal income in state over year, an index indicating the degree of personal freedom from arrest and incarceration in state, total population, and percentage net population growth rate in state. The authors have concluded that all of the variables except for the index of labor market freedom is statistically significant and agrees with some of the hypothesis put forward by previous literature (Cebula and Alexander 2020). However, the possible short coming of this research is endogeneity issue. Furthermore,

there could also be a collinearity issue in the regression. Precisely, it is plausible that total population growth correlates with the income level since the labor migration occurs based on urban area wages (Kennan and Walker 2011). Furthermore, the research does not answer about the effect of saving rate or about the effect of inequality. Thus, our research of focus is still unexplored in this literature. Other literature explores the factors leading people to homelessness (Early 2003), but the literatures are mostly on individual studies where individual survey data were used such as education, age, previous crime history, and so on. This individual case studies are in itself valuable but macroeconomic determinants are still left unanswered.

Theoretical works on this subject were done which examines individual decisions that lead people to homeless and the market conditions that potentially leads to homeless (O’Flaherty 2012). The research treats homelessness entries in the dynamic stochastic setting and considers the homelessness as part of some people’s life event. Paper finds that this way of looking at homelessness agrees with and also answers some of the empirical findings of the past. For instance, past homelessness is a significant predictor of future homelessness. It is because being a homeless is itself a shock to the asset and zero asset condition gives a rise to another future homelessness. Furthermore, homelessness is a stochastic process since otherwise it would be easy to predict homelessness which is empirically found to be false (O’Flaherty 2012). This paper highlights the importance of asset and consumption, which depletes asset (or future asset). Thus, in our paper, we would also consider the effect of saving in our analysis.

Data and Methodology

The primary question we are going to address in this paper is finding a macroeconomic factor that can explain the overall homeless population differences that exists between different

states. Aside from individual factors that lead people to homelessness, what macroeconomic factors can decide one to become homeless? We have reduced the potential variables to the following: saving rate, income inequality, and housing prices. First of all, the saving rate especially those that live in the lower income household is potentially a promising predictor of the homelessness. The hypothesis arises from the theoretical motivation. Using the dynamic stochastic differential equations, O’Flaherty has found that people are substantially more likely to become homeless if the asset level is low (O’Flaherty 2012). The reason simply is because of the fact that if the asset level is low, it is easy to deplete the capital through consumption if there was no income. Thus, you would become what he called “asset homeless” in which depletion of asset leads people to become homeless. In this paper, we would like to see if there is any connection between low saving level and local homeless ratio. According to our hypothesis, the lower saving rate should make people more susceptible to housing shocks or loss of jobs.

Other variable of interest is income inequality. The reason why we are interested in the local income inequality simply arises from the observation made in the study which modeled the migration (Kennan and Walker 2011). They found that mean income is the key driving factor of migration. Thus, higher average income attracts migrants to the local area. However, once they have moved to the area, they will have to find a job. If a local job pays about average income to all of those that migrated to the area, there would be no informal sector. However, we know this would contradict our everyday empirical observation. We would like to hypothesize that if there exists a higher pay disparity measured by Gini index, migrants are more likely to end up in the informal sector. The reason is that migrants, expecting to receive an average income, end up getting income that is far less than average if the income disparity is higher. The reason is that the mean income and median income is different and higher pay disparity leads to expectation

error as people make migration decision. Even if a migrant is expecting to receive average income, he/she is likely to receive median income or less; and thus, it creates the pool of people employed in the informal sector including potential homelessness (Harris and Todaro 1970). For this reason, we have included income dispersion in our analysis.

We have gathered state specific population data, average consumption, median housing price, Gini index, Average and median income, and employment rate data from US Census, and homeless data from USA Facts – all in 2021. The data observation amounts to 51. With the data above obtained, we have constructed local saving rate in the lower income distribution. We have derived two levels of saving rate. One is a saving rate of those that are at and above poverty level, denoted by s_{i1} , and the other one is a saving rate of those that are below poverty level, denoted by s_{i2} where i is an individual state respectively. The calculation is thus follows:

$$s_{i1} = \frac{\text{median}_{i,\text{income}} * \text{employment rate}_{i,\text{at and above poverty}} - \text{percapita consumption}_i}{\text{median}_{i,\text{income}}}$$

$$s_{i1} = \frac{\text{median}_{i,\text{income}} * \text{employment rate}_{i,\text{below poverty}} - \text{percapita consumption}_i}{\text{median}_{i,\text{income}}}$$

Furthermore, we have developed a measure of inequality via following calculation:

$$\text{inequality index} = \frac{\text{average}_{i,\text{income}} - \text{median}_{i,\text{income}}}{\text{median}_{i,\text{income}}}$$

Finally, to reflect the housing price in our analysis, we have divided housing price by income to obtain the percentage expenditure on housing as a variable. The obtained linear probability model incorporating savings, housing price, and inequality are as follows:

homeless ratio_i

$$= \beta_0 + \beta_1 \text{Saving rate} + \beta_2 \text{Housing per income} + \beta_3 \text{inequality index} \\ + \beta_4 \text{inequality index} * \text{Saving rate} + \varepsilon_i$$

In our analysis, we used inequality index as a measure of income inequality instead of Gini index. The reason is simply that inequality index we have developed captures the income dispersion and income expectation error that individual tends to make. Furthermore, we have found that our inequality index approximates the Gini index with adjusted R-squared of 0.934 (Appendix: Chart 1). Thus, we are going to use our own inequality index in our study. To estimate the homeless ratio of a given state, we have created four models. In our first two models, we did not include the interaction term in the regression – inequality index multiplied by saving rate. Moreover, for the model 1 we have run the regression using the saving rate s_{i1} (above poverty level saving rate), and in our second model, we included s_{i2} (below poverty level saving rate).

The results are summarized in Appendix: Table 1. The result tells us that both model 1 and 3 are statistically significant. Furthermore, the coefficient on the variable tells us the probability of becoming homeless since it is a linear probability model. Specifically, for both models, inequality index and saving index are both negative. However, for the model 2, we find that saving is statistically significant at 10% level, implying the importance of saving especially among the poverty level. Moreover, both model shows that housing expenditure as a percentage of income is a significant factor in predicting homeless ratio across states. This agrees with the previous study that also looked at the interstate differences (Cebula and Alexander 2004). The surprising result is that inequality is not a significant factor of homelessness at least in state level contrary to our hypothesis. For the third and fourth model, on the other hand, we have included

the interaction term and also run the regression using above poverty level saving and below poverty level saving for model 3 and model 4 respectively. Contrary to our first two models, the result shows us that interaction term has added a significant value in our regression. Coefficient on inequality index, saving, and housing per income are all positive, reflecting the fact that higher inequality leads to higher homelessness. The only change is that the saving coefficient became positive after the interaction term has been added. However, the marginal effect of adding one unit of saving is negative holding everything else constant. It is because the coefficient on the saving is positive but the coefficient on the interaction term is negative with higher magnitude, making the overall marginal effect negative. Thus, lower saving increases the likelihood of being homeless, which is consistent with our hypothesis. Moreover, if the saving rate is low, then the marginal effect of inequality becomes positive, meaning inequality makes it likely to create homeless only if there is a lower saving. On the other hand, higher saving implies that marginal effect of inequality becomes negative, meaning higher saving in below poverty level reduces homelessness regardless of the level of inequality. This reflects the importance of saving and inequality and, to address the homeless issue, both factors must be jointly considered.

Conclusion

In this paper, we have looked at the interstate differences in homeless population. We found that inequality, saving rate (especially of those below poverty rate), and housing per income are significant factors in explaining homeless ratio across different states. Our result is consistent in part with theories and previous literature and highlights the importance of inequality and saving rate and suggests that those factors must be considered jointly. However, although we have identified the association between inequality, saving, and housing price, we

have still yet to find the causal effect of those variables; thus, future research must be focused on the causality effect by using quasi-experimental design. Possible counterargument to our hypothesis is that perhaps, one can argue that homeless people – as determined by “need” and “generosity” (O’Flaherty 2019) – find city to be the most high-quality area that homeless can dwell (and thereby higher homeless ratio in city area), and those cities (or almost all cities) are simply happened to be characterized by inequality, expensive housing, and high poverty rate. With this line of reasoning, one can also argue that there is no relationship at all between inequality and lower saving rate in causing homelessness – in other words, the homelessness and inequality and expensive housing just happened to coincide in the city area, and one simply does not cause other. However, the result shows us that higher the inequality and the housing price, no matter which city, the homeless tend to exist more likely. Thus, our answer to such a challenge is that, while we admit that it is a possibility, cities characterized especially by higher inequality and lower saving are more likely than other cities to be associated with higher homelessness. This means that inequality and saving are still significant potential causes of homelessness and must be studied more carefully in future research.

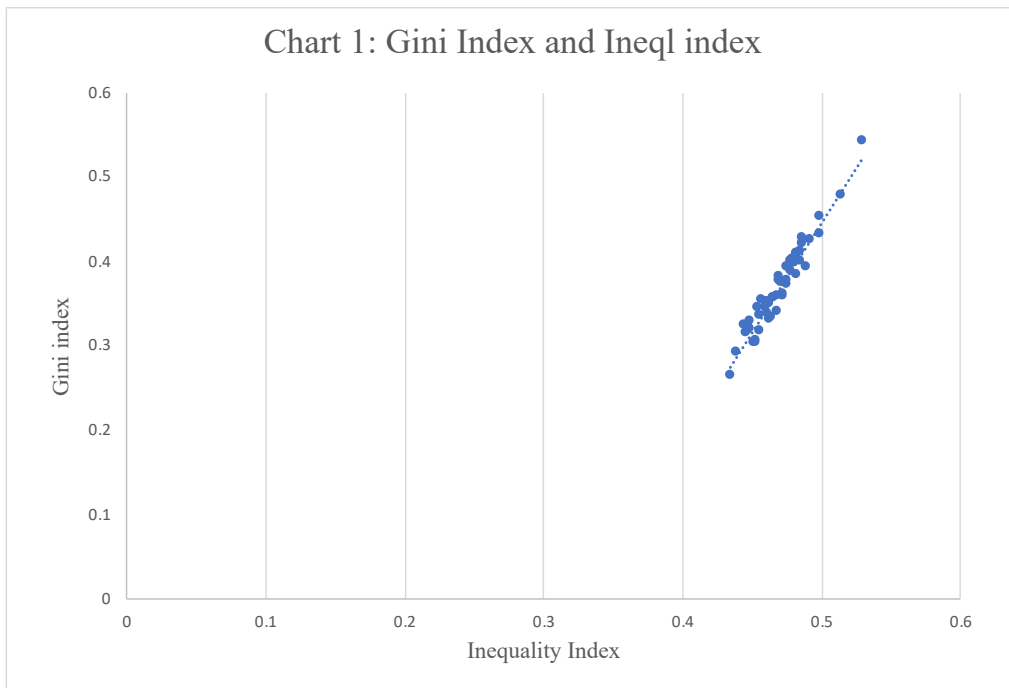
Appendix:

Table 1: An Example Table: Dependent Variable

	Model 1	Model 2	Model 3	Model 4
Inequality Index	-0.0032 (0.0044)	-0.0003 (0.0041)	0.0898*** (0.0165)	0.0236*** (0.0042)
Saving	-0.0028 (0.0038)	-0.0054* (0.0023)	0.0714*** (0.0141)	0.0403*** (0.0063)
Housing/Income	0.3772*** (0.0923)	0.3480*** (0.0853)	0.3929*** (0.0733)	0.4319*** (0.0590)
Saving*Inequality			-0.1986*** (0.0371)	-0.1182*** (0.0158)
Poverty level	Above	Below	Above	Below
Adj R ²	0.35	0.41	0.59	0.73
Model P-value	Less than 1%	Less than 1%	Less than 1%	Less than 1%

Notes: *: Significant at the 10% level; **: Significant at the 5% level; ***: Significant at the 1% level. Standard errors

listed in parentheses below the coefficients.



Works Cited

- Cebula, Richard J., and Gigi M. Alexander. "Economic and Noneconomic Factors Influencing Geographic Differentials in Homelessness: An Exploratory State-Level Analysis." *The American Journal of Economics and Sociology*, vol. 79, no. 2, 2020, pp. 511-540. doi: 10.1111/ajes.12320.
- Early, Dirk W. "The determinants of homelessness and the targeting of housing assistance." *Journal of Urban Economics* 55 (2004): 195-214. doi:10.1016/j.jue.2003.09.005.
- Harris, John R., and Michael P. Todaro. "Migration, Unemployment and Development: A Two-Sector Analysis." *The American Economic Review*, vol. 60, no. 1, 1970, pp. 126-42. *JSTOR*, <http://www.jstor.org/stable/1807860>.
- K. Corinth, D. S. Lucas, "When warm and cold don't mix: The implications of climate for the determinants of homelessness", in *Journal of Housing Economics*, vol. 41, pp. 45-56, 2018. <https://doi.org/10.1016/j.jhe.2018.01.001>
- K. Corinth, "The impact of permanent supportive housing on homeless populations", in *Journal of Housing Economics*, vol. 35, pp. 69-84, 2017. <https://doi.org/10.1016/j.jhe.2017.01.006>.
- Kennan, John, and James R. Walker. "The effect of expected income on individual migration decisions." *Econometrica*, vol. 79, no. 1, 2011, pp. 211-251. DOI: 10.3982/ECTA4657.
- Ministry of Health, Labour and Welfare (MHLW) website. (n.d.). Retrieved from https://www.mhlw.go.jp/stf/newpage_17922.html
- O'Flaherty, Brendan. "Homelessness research: A guide for economists (and friends)." *Journal of Housing Economics*, vol. 44, 2019, pp. 1-25. doi: 10.1016/j.jhe.2019.01.003.

O'Flaherty, B. (2012). Individual homelessness: Entries, exits, and policy. *Journal of Housing Economics*, 21(2), 77-100. ISSN 1051-1377. <https://doi.org/10.1016/j.jhe.2012.04.006>.

USA Facts. (n.d.). Homeless population. Retrieved from <https://usafacts.org/data/topics/people-society/poverty/public-housing/homeless-population/>

U.S. Census Bureau. (n.d.). Data.census.gov. Retrieved from <https://data.census.gov/>