UNIVERSITY OF CALIFORNIA,
IRVINE

Helping the Marginalized: An Empirical Analysis of Two Low-Income Groups

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Economics

by

Justin Jarvis

Dissertation Committee:
Professor Jan Brueckner, Chair
Professor Priya Ranjan
Professor Stergios Skaperdas

2015

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank Jan Brueckner for a great deal of support and many good ideas.

I would also like acknowledge that funding for the last chapter of my dissertation came from the multidisciplinary design program at UC Irvine.

CURRICULUM VITAE

Justin Jarvis

EDUCATION

Doctor of Philosophy in Economics — **2015**
University of California, Irvine — *Irvine, California*

Master of Arts in Economics — **2013**
University of California, Irvine — *Irvine, California*

Bachelor of Arts in Economics, with Distinction — **2007**
University of California, San Diego — *La Jolla, California*

RESEARCH EXPERIENCE

Ethnographic census of street homeless — **2012–2015**
University of California, Irvine — *Irvine, California*

Survey data collection, US – Mexico migrants — **2006–2007**
Tlacuitapa, Mexico

TEACHING EXPERIENCE

Instructor — **2013–2015**
Coast Community College District — *Costa Mesa, California*

ABSTRACT OF THE DISSERTATION

Helping the Marginalized: An Empirical Analysis of Two Low-Income Groups

By

Justin Jarvis

Doctor of Philosophy in Economics

University of California, Irvine, 2015

Professor Jan Brueckner, Chair

The three chapters in this dissertation analyze data from studies done with a focus on understanding marginalized people groups. The first analysis (in Chapter 1) looks at the lives of children in poor households in Ecuador, in particular by estimating the effects of the household-level determinants on educational achievement and on the incidence of child labor. The second two analyses (Chapters 2 and 3) attempt to understand the determinants of homelessness in Costa Mesa.

Chapter 1 finds that the Bono Desarrollo Humano in Ecuador has a positive effect on reducing child labor in the current generation, and perhaps secondary effects into subsequent generations. I also find that child labor and household educational attainment have a negative correlation, even after controlling for other observables.

Chapter 2 utilizes a novel dataset and to study homelessness and finds that the self-reported cause of homelessness is a determining factor. It also finds, as would be expected, that education and the presence of family nearby lessens the intensity of homelessness.

Chapter 3 investigates how a homeless individual's intensity of homelessness is predicted by the individual's discount rate and attitudes toward risk. The ex-ante hypothesis is that an increase in homeless intensity could be driven by an decrease in patience and/or an

increase in risk tolerance. The hypothesis is tested by using a novel application of weighted least squares to regress the individual's homeless intensity (HIM) on the risk and discount variables mentioned above, as well as on the clearly exogenous individual-level characteristics. Evidence is found to support the hypothesis, in that a decrease in patience is associated with an increase in the individual's HIM. Increasing levels of risk tolerance are associated with an increase in the HIM as well, and this effect appears to be significant even with the inclusion of observably risky behaviors.

The data for Chapter 1 was obtained from the Ministry of Social Development of the Government of Ecuador. The data for Chapters 2 and 3 come from two surveys the author conducted himself. The first took place in October of 2012 and the second occurred in April 2014.

I acknowledge financial support from a fellowship with the Multidisciplinary Design Program at the University of California, Irvine.

Chapter 1

Cash Transfers and Household Outcomes in Ecuador

1.1 Introduction

Child labor is a very real phenomenon in the current world. The International Labor Office (ILO) defines child labor as "work that deprives children of their childhood, their potential and their dignity, and that is harmful to physical and mental development."[1] In this paper, I take "child labor" to mean any child age 5-17 who is employed outside the home and receives a regular salary. According to the ILO, in the year 2008, there were 215 million children aged 5-17 worldwide who were engaged in child labor. This is 7 million fewer children than were laboring in 2004 (Diallo et al., 2008). The cause behind this decline is of great interest to policy makers, many of whom believe we have a moral obligation to children to reduce or eliminate completely child labor in this modern world. Another issue of interest is how to accelerate this reduction in child laboring. These topics motivate my empirical investigation.

[1] http://www.ilo.org/ipec/facts/lang-en/index.htm

It is a common goal of policy makers to desire to reduce child labor, and at the same time increase educational attainment. But the question frequently asked is, “What is the best method?” Cash transfers from the government to poor, laboring households are one current policy tool used to reduce child labor. The BDH program in Ecuador is one such program. However, it is necessary to show that this type of program brings about a reduction in child labor in the most efficient manner possible. The benefit is that a properly designed, targeted program can increase the household’s income and, the literature states, decrease the amount of child labor the household supplies to the market (Attanasio et al., 2006; Edmonds, 2006). The downside is that these programs can be prohibitively expensive. The government is, after all, paying large sums of money to portions of the population, in addition to administrative costs. The BDH program cost $60m USD to administer, which was paid for by a loan from the World Bank (World Bank, 2006). This does not not include the transfer payments, which themselves amount to almost 1% of the nation’s GDP each year (Schady and Araujo, 2006). It is therefore important to make sure that these very costly programs are achieving the results they have promised. Otherwise, the developing countries are needlessly indebting themselves.

In the context of this paper, the BDH is a cash transfer that some have claimed does reduce child labor by up to 40% (Edmonds and Schady, 2008). This number would imply an incredibly successful program in Ecuador which should be mirrored in low-income countries across the world. If, however, evidence presents itself that indicates the program was less successful, perhaps countries should take a different approach. Thus, it is important to investigate the plausibility of this magnitude, as well as the plausibility of cash transfer programs worldwide. In the case of social programs that are less successful than hoped for, it is important to evaluate other policy instruments that also reduce child labor, with the goal of finding and implementing the policy that is the least expensive on the margin. That is, the program that requires the fewest dollars per one unit reduction in child labor is the desired choice.

1.1.1 Bono Desarrollo Humano Program

Beginning in 2003, Ecuador began an unconditional cash transfer program that was targeted to the two poorest quintiles of the population. In accordance with this program, qualifying (those who were in the poorest two-fifths) households were paid $15 USD monthly, which represents about 10% of their total expenditures. Although originally intended as a conditional cash transfer, conditioned on child school enrollment much like the oft-studied Progresa program in Mexico, conditions were never attached to the money as the government did not have the ability to check that the households were in fact completing the conditions (Edmonds and Schady, 2008).

1.1.2 Data

The data used for this paper comes from the Ministry of Social Development of the Government of Ecuador, and is publicly available.[2]

From the beginning of the BDH program, 4 of the 24 provinces in Ecuador (Carchi, Cotopaxi, Tungurahua, and Imbabura) were selected for evaluation (see Figure 1.1). In each of these selected provinces, parishes were randomly drawn. In the selected parishes, qualifying households were randomly sorted into treatment (receive the BDH transfer) and control (do not receive) groups. An important difference to note from the Progresa program is that here the randomization occurs at the household level, and not at the village level. This helps prevent village level fixed effects from being picked up by the estimates.

In the selected parishes, all qualifying households, regardless of being in the treatment or control group, were administered baseline surveys from June to August of 2003. These were administered by the Catholic University of Ecuador, an organization separate from the

[2]http://www.registrosocial.gob.ec/rs/docs/

Figure 1.1: Provinces of Ecuador

government and thus less likely to induce recall bias in the subjects. Once these were completed, the treatment households began receiving the BDH money. A follow up survey was then administered between January and March of 2005. These surveys were nearly identical, and included modules on household composition, education, child labor, expenditures, and health. It is the data from these surveys that I use.

Attrition in the panel is very low, as 94.1% of the households present in the baseline survey were re-interviewed. There is no correlation between having attrited and receiving the BDH treatment. There is also no correlation between having attrited and any other observable characteristic of the survey, other than age (which happens because the second survey took place 18 months later).

I also constructed two rainfall deviation indicators to serve as a proxy for total household income. Using data from Ecuador's Direccion General de Aviacion Civil (DGAC, the Ecuadorean version of the FAA), I compute average monthly rainfall data from each weather station, using the data available (most weather stations had historical data for about the last 12 years). I then calculate the deviation from the mean rainfall for the 3 months immediately preceding the survey. I also construct a "shock" dummy variable which takes the value 1 if the rainfall in the 3 months preceding the survey was more than 15% away from the mean. Finally, using GPS data from DGAC and Google maps, I match each province to the nearest weather station. These shocks affect nearly 25% of the households.

I also construct a variable to measure average household educational achievement. I take the current educational achievement (that is, grade level) of each child in the household and then divide it by the maximum possible achievement for her age. I then average out this ratio among all children in the household. The result is a random variable that ranges from 0 to 1, with zero meaning no educational attainment and 1 meaning that every child aged 5-17 in the household has the maximum education available for their age. I then discretize this variable to be able to use it in the multivariate ordinal probit setup. Sorting the households by the educational achievement variable, I classify households below the first quartile as LOW, those within the interquartile range as MEDIUM, and finally, those above the third quartile as HIGH education households.

There exist take-up and leakage problems in the BDH data set. In the sample studied in this paper, only 96% of households who are assigned to the treatment group report receiving the payments. This may be for reasons of social stigma, or perhaps the trip to the bank is too costly. Additionally, of those assigned to the control group, 18% report receiving the BDH, though they should not. Apparently, at the beginning of the evaluation period, some administrators began giving the cash transfer to all qualifying households, and not just to the treatment group. By the time this error was discovered, it was decided unfair to stop the transfer payments to these "leaked" families. Thus, they continued receiving the transfer payments.

This problem, however, has a proposed solution (Edmonds and Schady, 2008). Because the administrators who "jumped the gun" appeared to have done so randomly (or at least not in any systematic manner; they were, after all, just uninformed), this leakage should not introduce any bias into the estimates. Rather, it should just affect the magnitudes of the reported estimates. Thus, one can instrument the (perhaps) non-random receive indicator with the random treatment indicator. Then, one can use the predicted (and still random) $\widehat{\text{receive}}$ as the treatment indicator in analysis. This is the solution I use throughout the paper whenever I refer to the treatment indicator. Intuitively, what this does is rescale the coefficients by the amount of leakage; in this case we are just rescaling the coefficients by dividing them by .96.

I restrict the sample to those households with children aged 5-17 in the follow up survey. I also restrict the sample to those households that have entries for all of the covariates (a necessary convention). This leaves 1,309 households in the sample, which represents 3,004 children aged 5-17.

1.1.3 Descriptive Statistics

Table 1.1 presents some descriptive statistics at baseline for the BDH data set. Column means are reported, with standard errors in parenthesis to facilitate comparison of the differences of the means.

The mean number of child laborers per household is reported. In keeping with the ILO definition, I define a child laborer as a child between 5 and 17 who works outside the home, not for a family business, and receives a regular wage. This mean is influenced by the number of households who have no child laborers in their household. In the BDH sample, only 301 of 1309 households report child labor. For those who do, the median number of child laborers is 1.

The values of the rest of the variables are typical for low-income households in a developing country. The educational achievement of the household head is low, the majority live in rural areas and the biweekly per capita numbers show us that each person lives on just greater than $2 per day, a level the World Bank defines as the minimum level for an "economic gradient to appear." (Ravallion et al., 2009)

It is easy to see that the randomization of the program was successful. A t-test for the difference in means between the treatment and control groups fails to find a difference, with the exception of the percent of households who are required to pay for the school bus (indicated by an asterisk). It is unclear why this is significantly different, but it should not jeopardize the reliability of the results.

In Table 1.2 I present the values of a few indicators from the follow up survey. The variables selected are those that intuitively would seem to be good candidates to change in the treatment households after receiving the BDH.[3] At first glance it appears that the mean number

[3]Sadly, the biweekly expenditure variable, which definitely should have changed, is not available in the follow-up survey

Table 1.1: Baseline Summary Statistics

	Treatment	Control	Overall
# child laborers/HH	0.21	0.23	0.22
	(0.02)	(0.02)	
Sum of child labor hours/HH/week	5.72	6.20	5.94
	(1.50)	(2.50)	
Average education/HH	0.83	0.81	0.82
	(0.10)	(0.11)	
# chore laborers/HH	1.82	1.79	1.81
	(1.16)	(1.12)	
Sum of chore hours/HH/week	17.76	17.19	17.46
	(16.92)	(18.40)	
HH pay for school bus	0.52*	0.42*	0.48
	(.05)	(.05)	
HH head's education (grade)	3.95	4.16	4.03
	(1.20)	(1.69)	
Family size	5.81	5.75	2.74
	(1.9)	(1.9)	
School Distance	27.13	26.2	25.5
	(18.89)	(19.6)	
Father is head of HH	0.84	0.86	0.85
	(0.36)	(0.34)	
Urban HH	0.45	0.48	0.46
	(0.05)	(0.05)	
Biweekly expenditure / capita (USD)	35.91	37.40	36.62
	(20.8)	(29.80)	
N	694	615	1309

*statistically significant difference

of child laborers increases from .21 to .23 upon receiving the cash transfer. Recall however, that this difference was already present in the baseline, and a t-test for the difference in means fails to reject the null hypothesis of equality. Furthermore, not a single variable in the treatment group varies significantly from its comparison group analogue. Thus, in the next section I fit a statistical model to disentangle the individual effects.

Table 1.2: Follow-Up Summary Statistics

	Treatment	Control	Overall
# child laborers/HH	0.23	0.21	0.22
Average education/HH	0.81	0.83	0.21
# chore laborers/HH	1.72	1.53	1.64
HH pay for bus	0.51	0.46	0.48
Family size	2.78	2.54	2.68

1.2 MVOP effects of the BDH program

In this section I update the current literature by using a multivariate ordinal probit (MVOP) model to estimate the effects that the BDH program has had on child labor and educational attainment at the household level. I argue that this model more properly fits the data. Previous research has been focused at the level of the individual child, and has used linear probability models (Schady and Araujo, 2006; Edmonds and Schady, 2008). The MVOP allows for non-monotonic effects across the sample space. Even though both of the mentioned papers use linear probability models, they discuss the fact that the effect of the BDH will not be linear across the whole parameter space, particularly in regards to income. That is, the BDH program will have a different effect on "richer" households vis-a-vis "poorer" households. This paper is the first attempt to quantify that numerically. With the MVOP model, the effect of all the covariates are allowed to be non-linear. Even if the effect of the variables of interest are found to be linear, the linear probability model is still most likely mis-specified if other covariates are found to have non-linear effects. This is the first reason

I argue that the BDH data is more appropriately suited to the MVOP.

The second reason is that it will allow me to estimate the covariance matrix of child labor and educational attainment. An original feature of this paper is that I am be able to see how child labor at the household level covaries with educational attainment, after controlling for the other covariates. I find negative covariances between the two variables, which means that I find them to be substitutes. This is consistent with previous BDH literature (Schady and Araujo, 2006), but different than data from a Bangladeshi study (Ravallion and Wodon, 2002). For these reasons I argue that the MVOP model is more appropriate to the data.

I first offer an intuitive explanation of how the model works (Section 1.2.1). I then explain the estimation methods (Section 1.2.2), specify the particular model I am fitting (Section 1.2.3), and offer the results (Section 1.2.4). Finally, I propose a test of the Luxury Axiom (Section 1.2.5).

1.2.1 Explanation of MVOP model

The model is easy to understand. A household works to maximize utility by allocating children to various tasks. We want to know how many child laborers the household supplies to the market: No child laborers, or perhaps one, two, or more? How many children are then allocated to schooling? I measure this by asking if the household is classified as having low, medium, or high educational achievement. And finally, I also ask how many children does the household dedicate to chores around the house? I believe this activity can affect the others mentioned, so is included in the analysis, and is modeled jointly. The number of children allocated to each task is the outcome variable. This can be modeled accurately with a multivariate ordinal probit model because the data has the following characteristics:

- The outcome variable is discrete. Households do not have the option to supply π

children, or $\frac{2}{3}$ children, to the labor market.

- The outcome variable is ordered. It makes sense to talk about two children being more than 1 child. This point also illustrates why a traditional technique, such as OLS, is inappropriate. The main assumption we use to model this household's decision is that the household is working to maximize its utility. Thus, if we observe four children, we know that the household experiences greater utility with four children than with three children. However, we do not know how much greater. This is because utility is ordinal, and not cardinal. We cannot make the assumption that household utility experiences an equal increase from having three vs two children in the labor market, as from eight vs seven children. Indeed, after supplying seven children to the labor market, the increase in household utility from number eight is likely to be very small.

- The outcome is multivariate. Our households don't supply children to the labor market only. They also supply children to school, as well as to do domestic chores. For every household we model, we will have three pieces of data - the number of children allocated to laboring, the number engaged in domestic chores, and the average household educational achievement. We need to model this as a system of equations, since we don't expect these allocation decisions to be independent of each other. Additionally, the household can easily dedicate 1 child to two or more areas. Thus, we would expect to see covariances between the outcome variables.

Furthermore, in the multivariate ordinal probit, we hypothesize that the covariates don't directly affect the outcome variable. Rather, the covariates affect a latent (that is, "unobserved) variable, and it is the level of this variable that affects the outcome variable. For each outcome of interest (in this case we have 3), we call the outcome variable y, the latent variable z, and the vector of covariates $\mathbf{X}$. Thus, the mathematical framework looks like this:

$$z_{1i} = \mathbf{X}_i\beta_1 + \varepsilon_{1i} \text{ for outcome 1 and,} \tag{1.1}$$

$$z_{2i} = \mathbf{X}_i\beta_2 + \varepsilon_{2i} \text{ for outcome 2 and,} \tag{1.2}$$

$$z_{3i} = \mathbf{X}_i\beta_3 + \varepsilon_{3i} \text{ for outcome 3 and,} \tag{1.3}$$

$$\varepsilon_i \sim N_3(0, \mathbf{\Omega}). \tag{1.4}$$

In this example, $\mathbf{\Omega}$ is a 3x3 symmetric, positive definite, covariance matrix. The estimate for this matrix tells us the amount that one outcome variable covaries with the others, after controlling for the other covariates. The observed y's are then generated from the following process, which discretizes the latent z's:

$$y_{1i} = j \ \textit{if} \ \gamma_{1,j-1} \leqslant z_{i1} < \gamma_{1,j} \text{ for outcome 1 and,} \tag{1.5}$$

$$y_{2i} = j \ \textit{if} \ \gamma_{2,j-1} \leqslant z_{2i} < \gamma_{2,j} \text{ for outcome 2 and,} \tag{1.6}$$

$$y_{3i} = j \ \textit{if} \ \gamma_{3,j-1} \leqslant z_{3i} < \gamma_{3,j} \text{ for outcome 3.} \tag{1.7}$$

where i = 1, 2, ..., N and is the index of observations.

$\gamma_{1,j}$, $\gamma_{2,j}$, and $\gamma_{3,j}$ are cut-points which determine the discretization of y_i, a vector of length three, which can take the values $y_{1i} \in \{1,\ 2,\ 3, ...\}$, and the same for $y_{2i},\ y_{3i}$. The index j is the index of the cut-points in each equation. This determines into how many "bins" the discretized y's will fall. Note that the number of cut-points need not be the same for each equation; in the framework above we have j_1 cut-points in equation 1 and j_2 cut-points in equation 2. The weak inequality can be equally applied to either side of the inequalities above. Another thing to note is that the equations need not have identical covariates, or even an identical number of covariates. The vectors β_1 and β_2 each have lengths k_1 and k_2, respectively, where k_1 is the index of covariates in equation 1.

1.2.2 Estimation Plan of Attack

The problem with MLE

The multivariate ordinal probit model is very difficult to estimate using the traditional Maximum Likelihood Estimation (MLE) approach. This is because using an MLE approach, the likelihood contribution of each individual is calculated, and the estimated beta is the vector of parameter values which maximizes the product of all the individual likelihood contributions. Formally,

$$\hat{\theta}_{MLE} = \underset{\theta}{argmax}\ f(y|\theta), \text{ where} \tag{1.8}$$

$$f(y|\theta) = \prod_{i=1}^{n} f(y_i|\theta). \tag{1.9}$$

Here, $f(y|\theta)$ is the product of all the likelihood contributions from all i households. In the univariate (only one y) case, this is just the area under the gaussian pdf with mean $\mathbf{X}_i'\beta$ and variance σ^2 , and between the two nearest cut-points (γ). Formally,

$$f(y|\beta,\gamma,\sigma) = \prod_{i=1}^{n} f(y_i|\beta,\gamma,\sigma), \text{ where} \tag{1.10}$$

$$f(y_i|\beta,\gamma,\sigma) = \prod_{j=1}^{J}\left[\Phi\left(\frac{\gamma_j - X_i'\beta}{\sigma}\right) - \Phi\left(\frac{\gamma_{j-1} - X_i'\beta}{\sigma}\right)\right]^{1(y_i=j)} \tag{1.11}$$

Here, I have replaced θ with the specific parameters (β, γ, σ) that I am trying to estimate. $1(y_i = j)$ is an indicator function that becomes 1 when $y_i = j$. Intuitively, this tells us that if y_i (e.g., the number of children) is equal to 3 for household i, we evaluate the area under the gaussian pdf (again, with mean $\mathbf{X}_i'\beta$ and variance σ^2) between cut-points 2 and 3. I have also used the notation $\Phi(a)$ to indicate the value of the standard normal CDF below a. Furthermore, it is necessary to divide $\gamma_j - \mathbf{X}_i'\beta$ by σ to standardize the value, which allows us to use the standard normal CDF, $\Phi(a)$.

It is precisely at this point that we have difficulty in the multivariate case. The parameter space now extends in more than one dimension; in fact, it has dimensionality equal to the number of outcomes we are measuring. Therefore, in order to calculate the likelihood contribution of each individual, we would need to evaluate the area under the multivariate gaussian PDF between 2 cut-points in every dimension. Using the trivariate case above, the household's likelihood contribution now looks like:

$$f(y_i|\beta,\gamma,\Omega) = \int_{\gamma_{1,j-1}}^{\gamma_{1,j}} \int_{\gamma_{2,j-1}}^{\gamma_{2,j}} \int_{\gamma_{3,j-1}}^{\gamma_{3,j}} [N_3(\nu_i|\mathbf{X}_i'\beta,\Omega)]\, d\nu_{3i} d\nu_{2i} d\nu_{1i}. \tag{1.12}$$

N_3 is the three dimensional gaussian pdf. The issue here is evaluating the integral in all three dimensions. This can become difficult rather quickly, especially as the dimensionality rises. And this is just for one household. Maximizing the product of N of these contributions requires more processing power than is available.

Strategy - Use Bayes' Rule

Using techniques from the Bayesian arsenal, we are able to estimate the parameters without overloading the computing power currently available. By realizing we have the data generating process from which the latent z_i is generated, we can simulate the vector z and make use of Bayes' Theorem, which tells us:

$$\pi(\beta, \gamma, \mathbf{\Omega}, z|y) \;=\; f(y|\beta, \gamma, \mathbf{\Omega}, z)f(\beta)f(\gamma)f(\mathbf{\Omega}). \tag{1.13}$$

In other words, the complete data posterior is equal to the likelihood function augmented with the latent (unobserved) z, times the prior distributions we have for β, γ, and Ω. This is helpful because if we sample from the complete data posterior, $\pi(\beta, \gamma, \mathbf{\Omega}, z|y)$ and ignore z, we are left with draws from $\pi(\beta, \gamma, \mathbf{\Omega}|y)$, which is our goal. The Bayesian point estimates for these parameters are just the mean of this distribution. Thus, we have achieved point estimates without having to evaluate the multidimensional integrals mentioned above. I will now look at some identification issues which, once accounted for, will enable us to write the complete data posterior from which we desire to sample.

Identification

Equation 1.11 above is unidentified in light of two problems. The first is, for any value of a constant, c, the following is true: $\frac{\gamma_j - X_i'\beta}{\sigma} = \frac{c\gamma_j - \mathbf{X}_i' c\beta}{c\sigma}$. Thus there is no way to know if the estimated values are really β or $c\beta$. This is usually overcome by setting $\sigma = 1$ in the univariate case, and in the multivariate case, writing $\mathbf{\Omega}$ in correlation form, with 1's on the diagonals. This is known as a scale restriction.

The second issue is that of location. Even after setting σ =1 (or arranging $\mathbf{\Omega}$ in correlation form), for any additional constant, g, $\gamma_j - \mathbf{X}_i'\beta \;=\; (\gamma_j + g) - (\mathbf{X}_i'\beta + g)$. In other words, we don't know if the estimated values are really γ or $\gamma + g$. This is generally overcome by setting $\gamma_1 = 0$. This also has the benefit of making estimation easier, as the vector γ now has $J - 1$ free (needing to be estimated) elements, instead of J. I set this restriction in addition to setting $\gamma_0 = -\infty$ and $\gamma_J = \infty$, which is not done for identification, but is required to encompass all values of $\mathbf{X}_i'\beta$, no matter how extreme, and is not a restriction itself.

Once the location problem is fixed by setting $\gamma_1 = 0$, there is another way to fix the scale problem without restricting σ (or $\mathbf{\Omega}$ in the multivariate case) as discussed above. This is accomplished by fixing another cut-point for each outcome variable. This is especially useful in the multivariate case, as it allows $\mathbf{\Omega}$ to be estimated freely, as a covariance matrix, instead of a correlation matrix.

For the remainder of this paper, I will use the location restriction $\gamma_1 = 0$ and $\gamma_2 = 1$, for each outcome variable, instead of a scale restriction. In the data used, each outcome variable has only 3 outcomes. This fortuitous result means that the vector of γ for each outcome does not need to be estimated. This simplifies the estimation process, as only β and $\mathbf{\Omega}$ are left to be estimated.

Estimation

Because of the identification scheme chosen (two fixed cutpoints for γ), we now have an identified model that can be estimated in the following approach, which is taken from Algorithm 6 in Jeliazkov et al. (2008). We first arrange the data so that each y_i contains all 3 of the outcomes for the individual i. We also arrange each $\mathbf{X_i}$ matrix in the following manner:

$$y_i = \begin{pmatrix} y_{1i} \\ y_{2i} \\ y_{3i} \end{pmatrix}, \ \mathbf{X}_i = \begin{pmatrix} (x_{1i,1} \ \ldots \ x_{1i,k_1}) & 0 & 0 \\ 0 & (x_{2i,1} \ \ldots \ x_{2i,k_2}) & 0 \\ 0 & 0 & (x_{3i,1} \ \ldots \ x_{3i,k_3}) \end{pmatrix}.$$

Once that is done, y_i and $\mathbf{X}_i$ are stacked in the manner normally used in seemingly unrelated regressions (SUR):

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \ \mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \vdots \\ \mathbf{X}_n \end{pmatrix}.$$

β, as usual, will have length equal to the width of $\mathbf{X}$, that is, $k = k_1 + k_2 + k_3$ (where k_1 is the number of covariates from equation 1). Given the preceding specifications, we have the following complete data posterior (an updated version of 1.13):

$$\begin{aligned} &\pi(\beta, \gamma, \mathbf{\Omega}, \mathbf{z}|\mathbf{y}) \propto \\ &\left\{ \prod_{i=1}^{n} \left[\prod_{m=1}^{3} 1\{\gamma_{m,j-1} \leqslant z_{i,m} < \gamma_{m,j}\} \right] N_3(z_i|\mathbf{X}_i'\beta, \mathbf{\Omega}) \right\} \\ &\times N_k(\beta|\beta_0, \mathbf{B}_0) \times W^{-1}(\mathbf{\Omega}|\mathbf{W}_0, v_0). \end{aligned} \tag{1.14}$$

The first term, in braces, is the product of the likelihood contributions of all n individuals making all 3 decisions. The second term is the prior on β , with prior mean β_0 (which has length k) and prior variance $\mathbf{B}_0$ (with dimensionality k x k). The final term is the prior distribution of $\mathbf{\Omega}$, which has an inverse Wishart distribution with inverse scale matrix $\mathbf{W}_0$ and degrees of freedom v_0 . Removing any restrictions on $\mathbf{\Omega}$ (as discussed in Section 1.2.2) allows us to sample it from the inverse Wishart distribution, which allows us to write the posterior above. Note that γ does not enter the posterior, since we previously fixed the cut points and is therefore not an estimable parameter. This simplifies the posterior distribution.

However, even though we can write the joint posterior distribution, we still need to sample from it. In order to get a point estimate, we must calculate the mean of this distribution. The easiest way to do this is to sample many draws from it, and then calculate the mean of those draws. Unfortunately, there is no direct way to sample from this distribution. Another method, the inverse-cdf method, is not helpful as we would still have to integrate this pdf to get the cdf, and then solve for the normalizing constant.

To get around this problem, we construct a Markov process which has this posterior as its invariant distribution. Recall that a process has the Markov property if each successive state depends only on the one immediately preceding it. In some cases, these processes will reach a point where every state is the same as the one before, i.e. the process has become stationary. We call the vector which describes this state the invariant distribution of the Markov Chain. In our analysis, we first construct a Markov chain whose invariant distribution is the posterior we want to sample from. We then successively run Monte Carlo simulations on this Markov chain, using the newly drawn values of the parameters as the conditioning parameters in subsequent runs. This will give us a vector of many draws from the posterior. To get to point estimates, we merely take the mean of these vectors, once the simulation has run enough times to be stationary (this is called *burn-in*).

The following algorithm (number 6 in Jeliazkov et al. (2008)) creates a Markov chain with the necessary invariant distribution:

1. Partition z_i , μ_i , $\mathbf{\Omega}$ as

$$z_i = \begin{pmatrix} z_{1,i} \\ z_{\neg 1,i} \end{pmatrix} , \ \mu_i = \begin{pmatrix} \mu_{1,i} \\ \mu_{\neg 1,i} \end{pmatrix} , \ \mathbf{\Omega} = \begin{pmatrix} \Omega_{1,1} & \mathbf{\Omega}_{1,\neg 1} \\ \mathbf{\Omega}_{\neg 1,1} & \mathbf{\Omega}_{\neg 1,\neg 1} \end{pmatrix}, \tag{1.15}$$

and define

$$\mu^*_{1,i} = \mu_{1,i} + \mathbf{\Omega}_{1,\neg 1} \times \mathbf{\Omega}^{-1}_{\neg 1,\neg 1} \times (z_{\neg 1,i} - \mu_{\neg 1,i})$$
$$\Omega^*_{1,1} = \Omega_{1,1} - \mathbf{\Omega}_{1,\neg 1} \times \mathbf{\Omega}^{-1}_{\neg 1,\neg 1} \times \mathbf{\Omega}_{1,\neg 1}. \tag{1.16}$$

2. Take one draw $z_{1i}|y, \beta, \gamma_1, \mathbf{\Omega}, z_{\neg 1,i} \sim TN_{\{\gamma_{1,j-1},\gamma_{1,j}\}}(\mu^*_{1,i}, \Omega^*_{1,1})$. $\mu^*_{1,i}$ and $\Omega^*_{1,1}$ are the conditional mean and variance, respectively, for z_{1i} given z_{2i} and z_{3i} . This is drawn from a truncated normal with cut-points $\gamma_{1,j-1}$ and $\gamma_{1,j}$, where $y_{1i} = j$, as discussed earlier. We then take one draw from z_{2i} and z_{3i} , just like we did with z_{1i} above. Then this is done for all $i = 1, \ldots, n$.

3. Sample from $\beta|\mathbf{z}, \mathbf{\Omega} \sim N_k(\beta^*, \mathbf{B}^*)$ where $\beta^* = \mathbf{B}^* \left(\mathbf{B}_0^{-1}\beta_0 + \sum_{i=1}^n \mathbf{X}_i' \Omega^{-1}\mathbf{z}_i\right)$, and $\mathbf{B}^* = \left(\mathbf{B}_0^{-1} + \sum_{i=1}^n \mathbf{X}_i'\Omega^{-1}\mathbf{X}_i\right)^{-1}$. Here, $\mathbf{B}_0$ and β_0 are the prior variance-covariance matrix, and the prior mean, respectively, for β .

4. Sample from $\mathbf{\Omega}|\mathbf{z}, \beta \sim W^{-1}\left(\mathbf{W}_0 + \sum_{i=1}^n (\mathbf{z}_i - \mathbf{X}_i\beta)(\mathbf{z}_i - \mathbf{X}_i\beta)', v_0 + n\right)$.

The above algorithm has several benefits. First, because we made use of cut-point restrictions exclusively to achieve identification, we are able to sample $\mathbf{\Omega}$ in an unrestricted manner from an inverse Wishart distribution. Furthermore, because each outcome variable y has only three response categories, the cut-points themselves (γ) do not need to be sampled. This makes the sampling easier by eliminating this step.

1.2.3 MVOP Model

I aim to answer the questions mentioned previously using this multivariate ordinal probit model. First, "How much of an effect does the BDH have on reducing child labor and increasing educational attainment?" For this, I look at the β coefficients from the model below. Also, “To what extent are child labor and child education substitutes?” To answer this I look at the correlation structure of all the left hand side variables.

The multivariate ordinal probit model I write is the following:

$$\begin{pmatrix} \#child.laborers_i \\ \#chores_i \\ ave.education_i \end{pmatrix} = \begin{pmatrix} \alpha_1 + \mathbf{X}_i'\beta_{1-6} \\ \alpha_2 + \mathbf{X}_i'\beta_{7-12} \\ \alpha_3 + \mathbf{X}_i'\beta_{13-18} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \\ \varepsilon_{3i} \end{pmatrix}$$

$$\begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \\ \varepsilon_{3i} \end{pmatrix} \sim N_3(0, \mathbf{\Omega}_{3x3}). \tag{1.17}$$

On the left hand side we have three discrete outcome variables. They are the number of children (aged 5-17) in household i who labor, who do household chores, and the average educational attainment for children in the house. Each of these can take the values $\{0,\ 1,\ 2\}$. For the $\#child.laborers$ and $\#chores$, this stands for 0, 1, 2 or more children in the category. For the *ave.education* variable, I have discretized the continuous variable that ran from 0 to 1 into 3 discrete units: LOW, MEDIUM, and HIGH, with the 1st and 3rd quartiles as cutoffs (see Section 2.2). This way each outcome variable has only 2 cut-points and I can use the algorithm detailed earlier in Section 1.2.2. The outcome variables here are discrete, ordered (2 children is certainly more than 1) and they are multivariate in the sense that they need to be estimated together. They need to be estimated together because each of these equations is related. This is because one single child can both engage in paid market work outside the home and do chores when she gets home. Thus, the variable $\#child.laborers_i$ is most certainly correlated with $\#chores_i$ and appropriately needs to be estimated at the same time in order to record the correlation (or variance-covariance) matrix.

Each equation has six, identical right-hand side covariates in $\mathbf{X}$. These are *receive*, *rain.shock*, *pay.bus*, *head.edu*, *hh.size*, and *sch.dist*, for each household i. The first is the "treatment" indicator, which is the predicted value of regressing the receive indicator on the treatment assignment indicator (see Section 2.2). Also included are dummies if the household received a rainfall shock and if it has to pay for school bus service, educational attainment (in years) of the head of household, the size of the household, and the school distance (in minutes). Because of the random assignment, it is not necessary to include control variables into the model. The variables included are all variables of interest. The covariates, because of the randomization, can be interpreted causally. In another specification, I included parish fixed effects to follow the current literature; however, these came out insignificant in every model. This makes sense because the randomization was done at the household level.

The model choice itself is not without challenges. Traditionally, the MVOP model is used to fit outcomes which are more like discrete "counts", with no limit to how high the number can be. In this case, however, the outcome variables $\#child.laborers_i$ and $\#chores_i$ are clearly limited to the number of children in the household. This is overcome by the specifying these variable to have only three bins: 0, 1, and 2 or more. As all the households have at least 2 children, every household could potentially be found in any of the three categories. This eliminates the "limit" problem.

There also exists the issue that having one child laborer in a four children family is very different from the nine-child family with one child laborer. For this reason, I included the variable *hh.size* which will capture the effect of having more children, on the probability of having child laborers and chore laborers. Because the MVOP model allows for non-linear effects across the sample space, I can account for the different effect a four-child family would have versus a larger family. The traditional linear probability model, used in other works on this topic, cannot account for the possibility of non-linear effects.

The model is fit using the algorithm detailed above. I used 10,000 MCMC draws, of which 1,000 were used as burn-in. I also use the coefficients from maximizing the ordinal probit likelihood in the three individual univariate cases as the prior for the beta coefficients (β_0). The prior for the variance-covariance matrix on β, which is denoted B_0, is the identity matrix. The prior for $\mathbf{\Omega}$ is also the identity matrix, and the prior degrees of freedom are 50.

1.2.4 Results

The mean of the posterior distribution for β, which are used as point estimates, are reported in Table 1.3 (standard error in parenthesis):

Table 1.3: Results from MVOP model

Coefficient	#child.laborers	#chores	ave.education
BDH Indicator	-0.28	0.06	0.31
	(0.08)	(0.07)	(0.06)
Rain Shock Indicator	0.28	-0.08	-0.06
	(0.07)	(0.07)	(0.05)
Pay for bus	-0.32	0.08	0.45
	(0.07)	(0.06)	(0.05)
Head Education	-0.06	0.01	0.10
	(0.01)	(0.01)	(0.01)
Family Size	0.02	0.18	-0.07
	(0.01)	(0.01)	(0.01)
School Distance	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)

Most of these coefficients have the direction and magnitude we would expect. Distance (in minutes) from school appears to have an effect on neither educational attainment nor the number of child laborers in a household. As family size goes up, it lessens average educational attainment, but slightly increases child labor, and greatly increases the number of children who do chores. This makes intuitive sense, as a larger family will have more children to assign to these tasks. Another strong effect is the head's level of education, which decreases child laboring but increases average household educational attainment. The coefficient on the dummy variable Pay for bus has a sign that is perhaps opposite to intuition; the family having to pay for a school bus means the family has higher educational attainment. This can be explained by understanding that many of these households have no way to transport their children to school. A school bus that costs money, while seen as a negative thing in developed countries, actually is a benefit; the children now have a ride to school. Perhaps most interestingly, the coefficient on the BDH indicator is significant in affecting the number of child laborers in a family and the average household education level. This result is consistent with another study analyzing the BDH Edmonds and Schady (2008). However, to really examine the influence of the BDH, we must look at the marginal effects.

These beta coefficients are only the estimated values of β from the model; they are not the marginal effects. They can give us an idea of the magnitude and direction of the effect, but not the actual effect. Due to the non-monotonicity the effect β has, as well as the way it influences lower levels of the ordered outcome variable differently than high levels, there is no way to write the marginal effect as a value. For example, increasing the household head's education by one year causes an increase in the probability of the family having no child laborers, and a decrease in the probability of the family having 2 or more child laborers. It is a function of many inputs, including the outcome level of interest, as well as the level of each covariate.

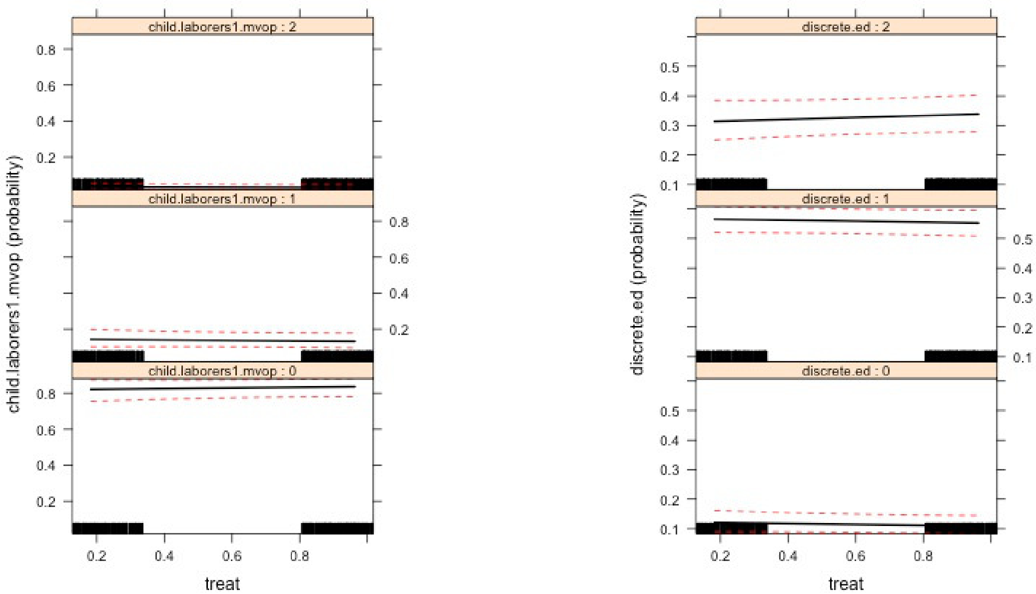
In this paper, I arrive at the marginal effect of a certain x, in this instance x_1, by the following process (similar to the one found in Section 5.1 of Jeliazkov et al. (2008)). First, I carefully select meaningful values for the values of the other covariates, $\mathbf{X}_{\neg 1}$. Generally, these are the column means, with the exception of the income variable, where I have chosen the first quartile. This is because in this paper I am interested in the effect the BDH has on poor households. I then compute $\Pr(y = j|x_1, \mathbf{X}_{\neg 1}, \beta)$. I repeat this step for every discrete value that x_1 can take. In the case of the treatment indicator, this is just 0 and 1. In the case of the household head's education, these are the integers from 0 to 16. For ease of explanation, I then plot these numbers in two dimensions (where the x-axis is the values of x_1, and the y-axis is the probabilities of being in each bin) and connect them with a smoothed curve. Finally, I draw this curve for all three values of j that y can take.

To analyze the effect of the BDH on the number of child laborers in a household, I consider the change in probability from when $\widehat{receive} = 0$ to $\widehat{receive} = 1$ (i.e., the family begins receiving the BDH payment). I do this for all three levels of the response variable.

In figure 1.2, I show the effect the BDH has on child labor and education. The left panel (Figure 1.2a) is the effect of the BDH on child labor. The bottom shows the probability of having 0 child laborers in the household, the middle having 1, and the top having 2 or more, with 95% confidence intervals. In the bottom left panel of figure 1.2a, I have calculated

$$\Pr(\#child.laborers_i = 0|\widehat{receive} = 0, \mathbf{X}_{i\neg 1}, \beta), \text{ and}$$
$$\Pr(\#child.laborers_i = 0|\widehat{receive} = 1, \mathbf{X}_{i\neg 1}, \beta),$$

and connected the probabilities with a line. $\mathbf{X}_{i\neg 1}$ is the vector of covariates, not including the treatment indicator, set at specific levels. As discussed, the levels have been chosen based upon the households we want to investigate. For the following figures, I report the marginal effects for an average poor family: they have not been victim of a rain shock, their household size is 5.7 and distance to school is 26.7 minutes (column averages). They do pay for the school bus, and their average monthly expenditure is $126.70 USD (this is the first quartile of the income column). This is therefore an average poor family who should, in theory, benefit from the program. We know, from the coefficients reported in Table 1.3, that receiving the BDH has the effect of decreasing the number of children laboring in each household. The bottom panel tells us that moving from an untreated household (left side) to treated household (right side), the probability of having 0 child laborers increases slightly, from about 81% to about 86%. Similarly, the middle panel tells us that moving from untreated to treated, the probability of having 1 child laborer in the house *decreases* from about 18% to about 14%.



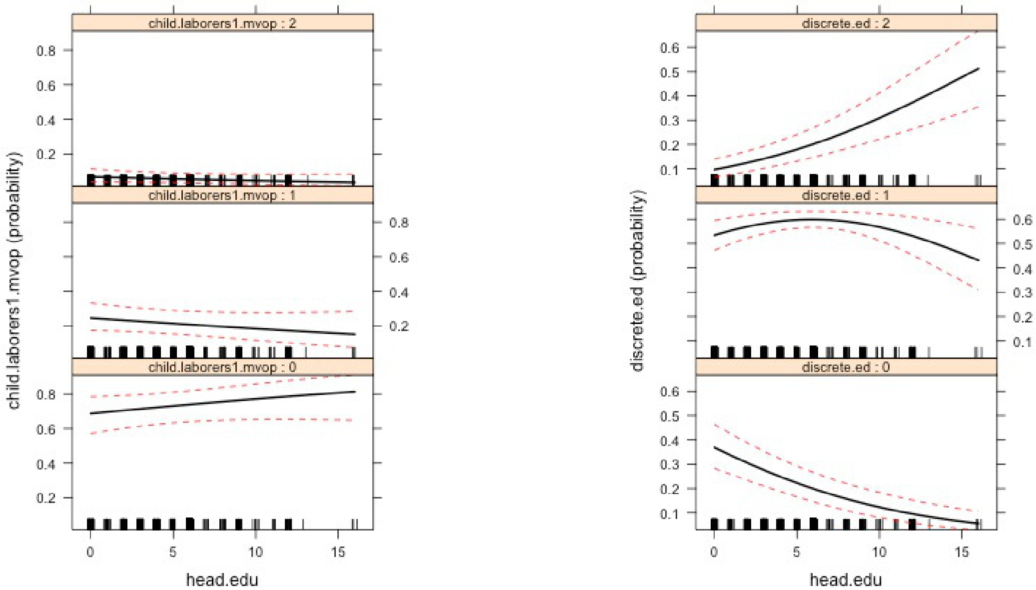
(a) on Child Labor

(b) on Ave Education

Figure 1.2: Marginal Effects of the BDH

Figure 1.2b shows the same thing, but for education. The bottom is the probability of being a household with low education, the middle is medium, and the top is high educational achievement. There exists a noticeable but slight upward trend in the upper panel; that is, moving from an untreated household to a treated household increases the probability that this household will have children with high levels of education.

In Figure 1.3 I show the effect that the head's education level has on the two outcome variables of interest, for the same hypothetical family. In addition, I set them as an untreated family (the graph looks the same if I were to show it for a treated family). Here we see much larger, and non-linear effects. Looking at the bottom panel of Figure 1.3a, we see that as the head's education rises, the probability of that household being a no-child laborer household increases. Looking at the upper panel of Figure 1.3b, we can see that as the head's education rises, the probability of being in the "high" educational achievement category rises quickly.



(a) on Child Labor

(b) on Ave Education

Figure 1.3: Marginal Effects of the Head's Education

Another interesting component is the correlation between the three outcome variables. The mean of the posterior distribution for $\mathbf{\Omega}$ gives us the following correlation matrix:

$$\mathbf{\Omega} = \begin{array}{c} \\ \#child.laborers \\ \#chores \\ ave.education \end{array} \begin{array}{c} \begin{array}{ccc} \#child.laborers & \#chores & ave.education \end{array} \\ \begin{pmatrix} 1 & 0.063 & -0.359 \\ 0.063 & 1 & 0.113 \\ -0.359 & 0.113 & 1 \end{pmatrix} \end{array}$$

This indicates medium- high levels of correlation between the latent z's which generate the observed y's. Most importantly, this seems to indicate that child labor and educational attainment are in fact substitutes (-0.359 correlation). Another way to think of this is that even after controlling for other variables, an increase in child labor causes a decrease in average household educational attainment. This is a much more rigorous analysis of the subject but still consistent with another study of the BDH program (Schady and Araujo, 2006). It is, however, inconsistent with a study in rural Bangladesh (Ravallion and Wodon, 2002). Conceptually, one can imagine that as child labor increases, the children have less time for school. There is also the possibility that by sending children to the child labor market, the family becomes accustomed to the children working, and thus values education less. Therefore, from a policy prescription standpoint, simply banning child labor may in fact increase educational achievement in households.

1.2.5 Test of the Luxury Axiom

In Basu and Van (1998)'s article, they propose a cause of child labor known as the luxury axiom. The luxury axiom states that child schooling is a luxury; once the household reaches the "luxury" cut-off (generally thought of as the perceived subsistence level), they can begin supplying children to the schools instead of to the labor market. Thus, a household only supplies children to the labor market if they are below a perceived subsistence threshold.

Explanation of Test

If the luxury axiom holds true, then households in general should cease providing children to the child labor market as their household expenditures rise. Thus, in the parlance of the multivariate ordinal probit model, the probability of the household being a no-child laborer household should increase as the household's total expenditure increases. This is the bottom panel in Figure 1.4. Furthermore, the probability of being a 1-child laborer household should decrease after a certain level of income. This is represented by the center panel.

Results

In Figure 1.4, I present the marginal effects of total expenditure on the probability of a household being a 0, 1, 2 or more child-laboring household. This is the same model as the one written in Section 1.2.3, with the addition of an extra covariate, total expenditure. I compute the effects in the same manner as was described in Section 1.2.4. I set the values of the other covariates to the values from that section as well. In the figure, we see that the lines are completely level; changing total expenditure has virtually no effect on the household's probability of being a 0, 1, or 2 or more child laborer household. In other words, as a household's expenditure (and also, wealth) increases, the household does not cease to provide labor to the child labor market. The wealthier households on the right side of the figure have the same probability as the poorer ones on the left of supplying one child to the labor market. Thus, I conclude that there is evidence against the luxury axiom in this data set.

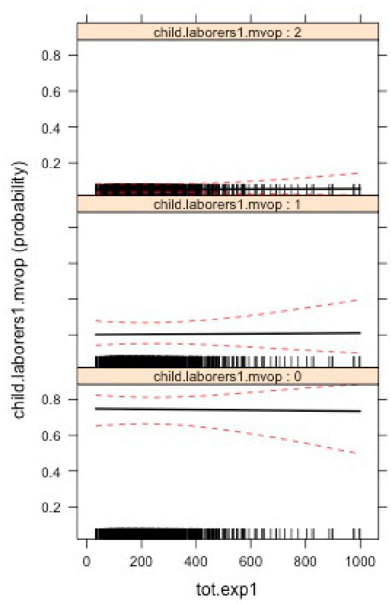


Figure 1.4: Effect of Income on Child Labor

1.3 Policy Implications and Conclusion

The findings in this investigation are full of implications for policy makers. The first is that overall, while the BDH did have an effect on the households, to the targeted (poor) households it only offered a small change in the probability of any one house being a child laborer household, versus a no-child laborer household. This is probably because the cash transfer was too small; $15 USD represents less than 10% of the average household's income. It is also much less than the foregone wage of a child laborer; the average child who works, works 40 hours a week and makes about $40 USD.

In order for the BDH to have a greater effect on child labor and educational attainment, the program will have to increase the size of the cash transfer. The program will probably have to be more targeted; currently it is given to 40% of the population. Because I have shown evidence against the luxury axiom, perhaps an income-based selection procedure is not appropriate here. From the variables I studied, not having a paid school bus was the biggest threat to increasing household educational attainment; it also was the biggest effect increasing child labor (see Table 1.2.4). Indeed, comparing conditional means, 35% of families with no paid bus service reported child labor, versus only 18% for those households with bus service. It seems more appropriate to perhaps select those households who are in the lowest two quintiles who also have no access to a paid school bus to receive the BDH cash transfer. Then with fewer households claiming the transfer, the amount per family could be increased. This then could move the budget constraint further for the affected households, thus perhaps effecting a bigger change in the variables of interest: child labor and educational attainment.

Chapter 2

Individual Determinants of Homelessness: A Descriptive Approach

2.1 Introduction

What is a homeless person? Why is this person without a home? Section 330(h)(4)(A) of the United States Public Health Service Act (PHSA) defines a homeless individual as "an individual who lacks housing (without regard to whether the individual is a member of a family), including an individual whose primary residence during the night is a supervised public or private facility (e.g., shelters) that provides temporary living accommodations, and an individual who is a resident in transitional housing."[1] It is estimated by the US Department of Housing and Urban Development (HUD) that on any night 643,000 people are homeless in the United States. During the 12 month period from October 2008 to

[1] The definition from the McKinney Homeless Assistance Act in 1987, Congress's first attempt at ameliorating the homeless problem, uses an almost identical definition.

September 2009, 1.56 million different people used an emergency homeless shelter. These numbers have grown significantly since the 1980's, but have decreased slightly since 2006 (HUD, 2009).

The numbers reflect the current two-pronged technique for counting the homeless in the United States. First, a point-in-time census is taken by a large number of counters during one particular night nationwide. This approach prevents double counting homeless people who may move to a new location every night. Second, researchers sum the number of different people who use one or more of the many forms of homeless shelters during a specified time period. Clearly, these numbers need not be the same, but they are used to give an indication of the number of people experiencing homelessness in the nation.

Finding and enumerating everyone who is experiencing homelessness on a particular night is a difficult task, particularly since most of these individuals do not wish to be found. Making this task even more difficult is the fact that no uniform definition of a homeless person exists. Commenting on this problem, O'Flaherty (1996, pg. 10) makes the following observation: "where you slept last night determined whether you were homeless. But...the sleeping places that qualified you as homeless were those where, it was thought, only vagrants slept." Different locales, even to this day, use different definitions for a homeless person (Jencks, 1995).

Research thus far on the determinants of homelessness falls into one of two groups. In the first group are studies that explore the topic from a macro level. Given the empirical challenges described in the previous section, such research has centered around linking city or state level variables such as median rent with the homelessness rate for the area (see, for example, Honig and Filer (1993); Quigley and Raphael (2001)). The general result found in these papers is that the city-level homeless rate rises with a decrease in median incomes or an increase in median rents.

In the other group are studies that have modeled homelessness on an individual level. See, for example, the recent Rolston et al. (2013) report, which found that the presence of a homelessness prevention program in New York City (known as Homebase) reduced the number of shelter entries by over 60 percent. However, Goodman et al. (2014), studying the same program, concluded that the average length of a homeless family's shelter stay was unaffected by participation in Homebase.

Several of the individual level studies have carried out homeless interviews to describe the nature of homelessness. Braga and Corno (2011) interviewed 62 percent of the street homeless individuals in Milan and found that the population is overwhelmingly male and middle aged (which this study also finds). Piliavin et al. (1993) found, in what is perhaps the most similar study to the one presented in this paper, that among a non-random sample of 330 homeless individuals in Minneapolis, the duration of homelessness is positively influenced by less consistent work histories and participation in the foster system, and is not influenced by symptoms of alcoholism.

This paper comes from the perspective of the latter group and intends to sharpen the methodology used in Piliavin et al. (1993) by using, for the first time, a methodology that mimics creation of a random sample of the homeless population. Previous studies have neglected to account for the bias introduced in the estimates when the most chronically homeless individuals are unintentionally oversampled. The adjustments described in Section 2.3 can be used to improve future attempts to utilize data on the homeless.

Using this novel data set, the paper answers the following question, "Given that an individual is homeless, what are the determinants of the intensity of homelessness?" The question is answered using survey data for the city of Costa Mesa, California. Homelessness intensity in this paper is defined as as the percentage of possible nights that an individual actually spends on the streets. I define "possible" nights as every night after the night of the individual's first homeless episode. Because homelessness is a transitory phenomenon, with individuals leaving and re-entering on any given night, it is perhaps more useful to think of homelessness as existing on a [0, 1] continuum, rather than being a dichotomous YES/NO variable. Therefore, this paper constructs a homeless intensity measure to augment (or replace) the common YES/NO measure of homelessness. It is useful to point out that the goal of this paper is *not* to describe the determinants of an individual falling into homelessness the first time (such as one would do by regressing the YES/NO homelessness measure on covariates in a probit model, for example).

The paper finds that the self-reported cause of homelessness and the level of education have large effects on the intensity of an individual's homelessness. The effects of other hypothesized factors, such as age, sex, mental health, and veteran status, appear to be less important.

Additionally, this paper evaluates an attempt by a coalition of churches to provide a useful intervention for homeless individuals. This intervention, known as the Check-in Center, is a location where homeless individuals can "check-in" their belongings during the day. The creators of the Check-in Center believe that the existence of such a location will enable the individuals to move around and to be productive during the day, instead of spending the whole day guarding their possessions. Homeless individuals unencumbered by possessions, it is thought, will be more capable of taking steps to end their homeless episodes. As hoped, this intervention appears to lessen the intensity of homelessness.

2.2 Data

In the fall of 2012, a group known as the Churches Consortium (with which the author of this paper is affiliated) launched an effort to characterize the homeless individuals living in Costa Mesa. To do so, data were sought from every single homeless person in the area. While there exists no uniform definition as to who qualifies as a homeless individual, as discussed earlier, the PHSA definition mentioned in Section 2.1 was followed as closely as possible. Therefore, the decision was made to proceed with a census in a two-pronged manner similar to that utilized by HUD.

First, the Consortium directed a corps of volunteers to systematically search the city of Costa Mesa from 5-8 am one morning in November. The volunteers were instructed not to approach, but rather count, each homeless individual. That morning, the volunteers counted approximately 150 people.

Next, the Consortium mobilized volunteers in an attempt to administer a detailed demographic survey (hereinafter referred to as the CM Assessment) to every single person experiencing street homelessness in the city. The volunteers first surveyed the individuals who were staying in homeless shelters. Next, they approached the soup kitchens. Finally, the volunteers moved to the streets to interview any last individuals who had not yet been surveyed. The survey was concluded after 1 week when persistent searching did not turn up any additional potential interviewees.

All of the respondents were asked if they were homeless; if they answered in the negative, their information was recorded but excluded from the data set. An incentive system was put into place to induce participation: all respondents were given the opportunity to participate in a party with their interviewers and community members at a local church at the conclusion of the survey period. With this structure in place, a very small number($< 5\%$) of potential respondents refused to be surveyed. Approximately 180 people are included in the CM Assessment. Using the probability weighting scheme discussed in Section 2.3.3, this number implies a total population of approximately 300 different individuals, not all of whom are homeless on any given night.

The census method described above systematically excluded individuals living in motels and/or with friends. Essentially, anyone not on the street was excluded. However, according to the PHSA definition, these people are also homeless. Therefore, in the survey portion, every effort was made to include this type of individual (by interviewing at soup kitchens, homeless gathering locations, etc.). This approach found approximately 40 people who did not sleep in an unsheltered location, individuals who would likely not be counted in the census (because they would not have been seen), but who came to service locations and thus were included in the CM Assessment. This is likely an explanation for the difference in the number of individuals counted in the census (150) versus the number counted in the CM Assessment (180).

Finally, using a number of "residential belonging" questions, it was determined which of the sampled individuals were Costa Mesa residents. Only these individuals were used for the purposes of estimating the model.

2.2.1 Descriptive Statistics

In Table 2.1, various population estimates from Costa Mesa (CM), produced by this study, are shown in comparison to broader averages. The CM averages here have been adjusted by the weighting process explained in more detail in Section 2.3.3. Among the averages in the table, Costa Mesa has a much higher percentage of white homeless individuals than is the norm. However, this is probably not a faulty finding but rather a real characteristic; the Orange County Continuum of Care (of which Costa Mesa is a part) estimates that 66.4% of its homeless are white (County of Orange, 2013, p. 10).

The broad demographic characteristics shown in Table 2.1 comprise the majority of what is known about individuals experiencing homelessness. While these characteristics are in fact recorded with a certain amount of accuracy, given the constraints, few other details are

[2] Bi-annual count done by the Los Angeles Homeless Services Authority.

Table 2.1: Comparison- Nationwide (HUD), Los Angeles, Costa Mesa homeless

Characteristic	HUD 2009 %	LAHSA[2] 2013 %	Costa Mesa %
Male	63.7	72.8	72.5
Female	36.3	27.2	27.5
White	38.1	36.6	72.0
Latin@	11.6	21.8	10.6
African American	38.7	38.0	3.4
Other Single Race	4.7	2.5	9.4
Multiple Races	7.0	1.1	4.5
Under Age 18	22.2	9.0	12.5
18 to 30	22.3	X	23.6
31 to 50	38.3	X	36.1
51 to 61	14.4	X	22.7
62 and over	2.8	8.0	5.1
Veteran	11.1	11.6	13.0
Disabled	37.8	18.2*	18.8*
Chronic Homeless	27.0	25.3	27.2**

* Only includes physical disabilities **more than 4 episodes in last 3 years

known on an individual level. One of the strengths of the novel data set created with the CM Assessment is the rich detail collected from each person as part of the survey process.

Both the sample statistics and the inferred population parameters (which come from weighting) are presented in Table 2.2. The correct way of interpreting the weighted and unweighted summary statistics is as follows. The inferred population parameter adjusts the sample statistic by realizing that some individuals in the population, those with low intensities of homelessness, have a lower probability of being interviewed, and are thus likely to be underrepresented in the sample. The ratio of total days homeless to total number of days since first homeless episode (essentially, percentage of life spent on street since one first became homeless) is used as the probability of inclusion. Therefore, the adjustment from sample statistic to inferred population parameter intends to capture the characteristics of the individuals who "weren't there," who may be homeless at some point but were housed when the sample was collected.

The sample average for age is 47.7, whereas the inferred population parameter is 46.1. We

Table 2.2: Summary Statistics

	Inferred Pop Mean	Sample Average	N
Age	46.1	47.7	175
Age at first episode	40.2	39.7	75
Total yrs hmls	5.3	6.0	170
Times hmls	2.9	3.8	70
% Adult life hmls	23.4	27.1	166
First time hmls	0.49	0.49	178
Male	0.72	0.71	181
Brain injury	0.32	0.30	174
Invol. mental hold	0.18	0.24	160
Inject RX	0.12	0.19	182
Drug treatment	0.20	0.27	182
Foster	0.12	0.09	180
Ever-incarcerated	0.69	0.73	182
Probation	0.17	0.16	167
VA status	0.13	0.16	178
Combat	0.03	0.06	182
Employed	0.10	0.15	182
Rented/owned in CM	0.49	0.53	176
Crime victim	0.41	0.41	180
Feel safe	0.70	0.72	175
Education (in %)			
K-8	2.5	3.1	
Some High School	15.1	17.9	
GED	3.9	5.6	
HS Graduate	19.1	19.0	
Some College	44.5	39.7	
College Grad	13.7	12.3	
Other	1.1	1.7	
Total	100	100	179

can take this to mean that our sample (which oversampled intensely homeless individuals) overrepresented older individuals; the younger individuals "weren't there" to be sampled because they have a lower intensity of homelessness. This therefore tells us something interesting: that among the homeless, being older tends to be positively correlated with more intense homelessness. This same phenomenon, which can be seen anytime the inferred population parameter is less than the sample average, occurs with the total years homeless, percent of adult life homeless, involuntary mental hold status,[3] injection drug usage, drug treatment, ever-incarcerated, veteran, and higher-education variables. This makes sense and should prompt the following note of caution: if inferences are made about the total homeless population based upon the individuals observed to be homeless at any one time, the homeless population will seem more mentally unstable, more drug-abusing, more incarcerated, less educated, and overall more intensely homeless than it, in fact, really is.

The means for brain injury and foster care appear to move in the "wrong" direction once weighted. The inferred population parameter for the proportion with a brain injury is .32, which is greater than the sample average. Using the same reasoning as used earlier in this section, but in the opposite direction, one might conclude that having a brain injury is *inversely* correlated with homelessness intensity. However, this is likely to not be the case and is probably an effect of sampling randomness.

2.2.2 Sleeping Location Last Night

In an effort to help the reader better understand this unfamiliar sector of the population, Table 2.3 presents the locations where the respondents reported sleeping last night.[4] This table presents only counts from the sample and therefore is not weighted, as the goal is not to

[3] California Welfare and Institutions Code, section 5150, allows for a 72 hour involuntary hold for people who are suspected of endangering themselves or others due to an untreated mental illness.

[4] The "feel safe" measure from Table 2.2 is the percentage of people who responded that they felt safe in the location where they slept last night.

make inferences about the whole population. Those places considered to be the "traditional" sleeping locations of individuals experiencing homelessness appear to be used frequently by the individuals studied in the CM Assessment. The plurality of the individuals slept directly on the street or sidewalk during the previous night. The second most frequently reported location was inside a vehicle. As has been reported by various sources (see, for example National Radio Project (2009)), the surge in homelessness which began in the 1980's saw many of the homeless individuals begin to live in their own vehicles. Other typical sleeping places. such as commercial buildings, parks, churches, and beaches, are represented in this sample as well.

Table 2.3: Slept Last Night (weighted)

Location	**Percent**
Street/Sidewalk	25.6
Friend/Family house	13.9
Vehicle	9.8
Church	8.7
Motel	7.5
In/around commercial building	6.7
Park	4.7
Beach	3.5
Treatment center	2.3
Other/unknown	17.4
Total	100

2.2.3 Self-Reported Causes of Homelessness, and Health

People become homeless due to various causes, and the paper investigates whether the self-reported cause has an effect on the intensity of homelessness. Table 2.4 lists the self-reported causes of homelessness and the relative frequency of occurrence. The interviewees responded with many different posited causes which were categorized into seven main groups.[5] Many individuals (37.5 percent) report that they became homeless because of a job loss or other

[5] The exhaustive list is also interesting, and is available upon request.

similar economic problems. An interesting statistic is that more than 10 percent of the interviewees explained that they became homeless due to what is labelled here as “Choice”; that is, they had a housing option but instead selected homelessness.

Table 2.4: Self-Reported Cause of Homelessness (weighted)

Location	**Percent**
Job Loss/Econ Problems	37.5
Relationship Discord	18.0
Choice	10.4
Loss of Home/Eviction	7.5
Alcohol/Drug Abuse	6.9
Injury/Health Problems	6.4
Other	13.2
Total	100

Table 2.5 reports the distribution of health among the interviewees. A large proportion, slightly greater than 20 percent, of the individuals report being disabled. However, the majority of individuals, about 70 percent, report being in at least fair condition.

Table 2.5: Self-Reported Overall Health (weighted)

Condition	**Percent**
Disabled	23.9
Very Poor	2.1
Poor	5.5
Fair	25.4
Good	43.1
Total	100

2.3 Sampling Methodology

Unfortunately, the sampling of Costa Mesa homeless individuals is subject to (at least) three main biases:

- Bias due to migration after losing housing

- Bias due to selection of convenient subjects
- Bias due to the transitory nature of homelessness

I attempt to control for each of these biases as explained in the following sections. The result is that the estimates presented in this paper can be treated as if they come from a random sample of the homeless population.

2.3.1 Migration Bias

If many Costa Mesa homeless individuals left the city after becoming homeless, then bias would be introduced into the analysis. It is likely that the migrants' location decision would be correlated with an unobservable, such as mental aptitude, that also influences homeless intensity. The result would be bias in the estimated coefficients of the covariates determining the intensity of homelessness.

It is therefore important to note the selection of Costa Mesa as the survey site. Costa Mesa can best be thought of as a "net receptor" of homeless individuals. Costa Mesa's mix of climate and services make it unlikely that a homeless individual would migrate away once becoming homeless. Indeed, one of Costa Mesa's pressing concerns (as outlined in the bulletin City of Costa Mesa 2012) is the prevention of non-Costa Mesa homeless taking up residence in Costa Mesa while simultaneously providing services to only those homeless who are Costa Mesa residents. Furthermore, nearly 80 percent of the interviewees identified Costa Mesa as their primary residence before becoming homeless. While this does not, in fact prove the contrapositive (that most Costa Mesa people who become homeless stay in Costa Mesa), it lets the reader understand the nature of migration flows among the homeless. Due to the aforementioned reasoning, by interviewing all the homeless in Costa Mesa and removing from them the non-Costa Mesa residents, we can arrive at a sample that is (largely) free of migration bias.

2.3.2 Selection Bias

Due to the difficulty of sampling this population, convenience sampling is very often used. For example, much of the individual-level data reported by HUD is reported to them by care agencies who sample those whom they can, with the count then scaled up to match the size of the whole population. This is known as snowball convenience sampling, in which the interviewee leads the team of researchers to another potential interviewee, who leads them to another, and so on. The discriminating reader can understand why this would lead to selection bias, as it is likely that the selection of interviewees in this scheme is correlated with an unobservable characteristic that might also be correlated with the outcome variable, intensity of homelessness. This bias, while at times unavoidable in the world of homeless research, is overcome in this study due to the sampling scheme: the (almost) simultaneous surveying of (almost) all homeless individuals in Costa Mesa.

By surveying every individual on the street at a given time, the selection biases inherent in the other methods are eliminated. This approach was only possible because Costa Mesa is a relatively small city ($\approx$16 mi^2) with a relatively small street homeless population. Furthermore, if a homeless individual doesn't want to be found, it is unlikely that he or she will be found. But by waiting at the common service locations (soup kitchens, etc.), the researchers in this study were able to interview the "recalcitrant" types when they came to replenish supplies. Thus, due to the unique characteristics of Costa Mesa, it is likely that the survey sample did not suffer from selection bias.

2.3.3 Transitory Bias

A random sample of this population, as is desired in most survey studies, is not easily obtained due to the fact the entire homeless population is not always available to be sampled. Thus, an alternate method was utilized and is described in detail in the following paragraphs.

Suppose that instead of randomly selecting individuals, a day was randomly selected and every homeless individual on the street was interviewed, across the whole city (call this the "census" method as described in Section 3.2.2). The results would comprise a sample of individuals, and sample statistics estimated from this sampling method would not be biased in the same way that estimates from convenience sampling are. However, in this census sampling method, the sample statistics are not consistent estimators of the population parameters, when the population is understood to be all individuals who are (or will be) homeless in a specified time period. This outcome is due to the fact that not all individuals are sampled with the same selection probability. The more intensely homeless individuals are more likely to be chosen, because they are on the street for a greater percentage of the time. For greater clarity, consider the following simplified example.

Suppose that in a finite population $U = \{1, \ldots, N\}$ there are two types of individuals, intensely- and mildly-homeless individuals, with N_1 intensely-homeless individuals, and $N_0 = (N - N_1)$ mildly-homeless individuals (recall, as mentioned in Section 2.1, that intensity refers to the percentage of a given time period spent sleeping on the streets). Set $y_i = 1$ for those in the intensely-homeless group, and $y_i = 0$ for those in the mildly-homeless group. The probability of being included in the sample on a random day is different for the two groups, because someone who is more intensely homeless is on the street more often, and more likely to be interviewed. Formally, $P(i \in S \mid y_i = 1) = \pi_1$ and $P(i \in S \mid y_i = 0) = \pi_0$, with $\pi_1 > \pi_0$, where S defines the chosen sample. $S = \{1, \ldots, n\}$.

Further suppose that we are interested in using S to make inferences about the whole population, and the parameter of interest is the proportion of individuals who are intensely homeless, that is, the average value of y in the population: $H = (N_1/N)$. Call the mean of the y_i's in the sample $\hat{H}$. By Bayes theorem, we can see that

$$\hat{H} = P(y_j = 1 \mid j \in S) = \frac{\pi_1 \times H}{\pi_1 \times H + \pi_0 \times (1 - H)}. \tag{2.1}$$

In this manner, it can be clearly seen that $\hat{H} = H$ only when $\pi_1 = \pi_0$. Pfeffermann (1996) explains that the solution to this difficulty is to weight by the inverse of the selection probability π. Define, generally, $w_i = 1/\pi_i$. The basic idea is that w_i is the number of "homeless persons" that person i represents. Thus, someone who is certain to be chosen will only represent himself, whereas someone with a small sample inclusion probability (but who was still included) must represent all of the other individuals with similarly small inclusion probabilities who were not included. Conveniently, it can be shown that $E_D(\sum_{i \in S} w_i) = N$, where $E_D(.)$ indicates the design expectation over all the possible samples that could be generated using the sampling scheme. In essence, taking the expectation over the design means that the statistic would asymptotically converge to this number if the sample were drawn an infinite number of times, and the statistic were calculated each time. This differs from the usual asymptotic justification of an estimator in that the sample size is not going to infinity; rather the number of samples chosen from a finite population is going to infinity (see Pfeffermann (1996); Carrington et al. (2000) for further explanation).

Thus, the estimator

$$\frac{1}{N} \sum_{i \in S} y_i w_i \tag{2.2}$$

is design-unbiased for H, and should be used instead of $\hat{H}$. However, since N is not known,

$$\frac{1}{\sum_{i \in S} w_i} \sum_{i \in S} y_i w_i \tag{2.3}$$

is the feasible design-unbiased estimator of H and is used in computing the statistics in Table 2.2.[6]

In similar fashion, Wooldridge (1999) shows that the presence of varying inclusion probabilities means that the usual OLS estimates from a regression carried out on the sample will be inconsistent because they assume equal probability of selection.[7] However, Wooldridge finds that weighting the criterion function (in the OLS case, the sum of squares) by the sample inclusion probability restores consistency. Weighting the criterion function by w_i is acheived by simply rewriting the population regression line so that:

$$\sqrt{\boldsymbol{W}} y = \sqrt{\boldsymbol{W}} \boldsymbol{X} \boldsymbol{\beta} + \sqrt{\boldsymbol{W}} \boldsymbol{\varepsilon}, \tag{2.4}$$

where $\boldsymbol{W}$ is an nxn matrix with w_i on the ith diagonal. From here, it is easy to see that the OLS estimator must be replaced by a weighted least squares estimator, $\hat{\boldsymbol{B}}_{\boldsymbol{wls}} = (\boldsymbol{X'WX})^{-\boldsymbol{1}}(\boldsymbol{X'WY})$.

However, the standard errors in this case are not the standard errors generated by the usual weighted least squares estimator that is used more frequently to deal with heteroskedasticity, as here we weight by the inverse of the root of the inclusion probability, instead of the inverse of the root of the variance. Kott (1991) provides a linearization estimator that is nearly unbiased for the variance of $\hat{\boldsymbol{B}}_{\boldsymbol{wls}}$, and this correction is used in this paper.

[6] In Table 2.2, I include both sample means (à la $\hat{H}$) and the estimates of the population parameters for comparison purposes.

[7] Solon et al. (2013) clarify, explaining that when the sampling probabilities are functions of the independent variables, OLS estimates are still unbiased. They explain that it is only when the inclusion probabilities are correlated with the outcome variable that OLS is biased. This latter case is the situation in this paper.

For the aforementioned reasons, the weighted least squares estimator is the appropriate design-unbiased estimator of the behavioral parameters and will be used in this paper. The ratio of total days spent homeless to total days since the first homeless episode (basically, the percentage of days spent on the streets since the individual first became homeless) is used as the sample inclusion probability π_i.[8]

Furthermore, weighting of the regression line gives us the coefficients of the effect on the overall possibly homeless population that we are implicitly seeking without realizing it. An example: suppose a policy-maker pursues policy X and desires to determine the effect of this policy on the"homeless." Further suppose that this person runs an unweighted regression to determine the effect of this policy and finds that it increases the probability of homelessness among those who are currently on the street (since the regression is unweighted, only claims about this smaller group can be made).

However, what is the negative effect on individuals in the greater homeless population who are currently housed? Not only will the HIM of the individuals currently on the street increase, but also the HIM of this greater population. And when the HIM of these currently housed individuals increases, they will end up on the street as well in the future. This is unaccounted for. When a policy maker desires to evaluate policy X, she wants to know the effect on all possible homeless individuals, not just those who are currently unhoused. The unweighted regression's results are biased.

Even worse, it is unknown in which direction the coefficients are biased. By imagining a scenario with heterogenous β's, in which the very intensely homeless are affected differentially than those who are currently housed, it can be seen that the direction of the bias thus depends on this (unaccounted for) difference between the groups.

Therefore, any policy maker objective function implicitly specifies the overall possibly home-

[8]This is true with one caveat: The probabilities were truncated to fall between five and 100 percent, inclusive. Probabilities that fell outside this interval were replaced with the nearest endpoint.

less population. That is to say, when a policy maker tries to reduce the numbers of homeless individuals on his or her streets, the desired reduction consists not only of the currently homeless individuals, but also a reduction in the number of possibly homeless individuals in the future.

Thus, by the arguments presented in the previous sections, the sample can be treated as a random sample of the homeless population of Costa Mesa, and as such the estimates are valid for population inference.

2.4 Homelessness Intensity Measure

In this section, one of this paper's contributions to the literature on homelessness is described: the concept of the intensity of an individual's homelessness. This paper utilizes one simple measure: The ratio of total days actually homeless (nights spent on street) to total potential days homeless (nights since first homeless episode began). Numerically, this is just the w_i weight mentioned in the previous section. As a result, regression observations end up being weighted by the value of the dependent variable.

The Homelessness Intensity Measure (HIM) is constructed by asking the interviewees the date they became homeless the first time. As the homeless population can be subject to recall problems, the interviewers were specifically trained to help the individuals with this question. The interviewers helped the homeless individuals recall the exact date by piecing together more easily recalled information, such as their age or the ages of their children at the time of their first homeless event. Then, the interviewer double-checked the date by asking something similar to, “So, you became homeless before Thanksgiving (or Christmas, or their birthday, etc.) in Year XX?” This procedure helped the individuals begin thinking about their homeless history in preparation for their next task, which was to recall specific details about the length of each of their homeless episodes. This detailed information was then aggregated by the interviewer and the total number of nights spent homeless was recorded. Both of these variables were recorded at the "month" detail level for ease of elicitation; if an individual was homeless for even one night in a month, the interviewer marked this person as being homeless the entire month when recording homelessness.

Determining the HIM using this two-step strategy also allowed the data to be checked for accuracy *ex-post*. If the numerator (total nights on street) was greater than the denominator (nights since first homeless episode), then this individual's HIM was marked as potentially incorrect. The regression was then run with and without the potentially incorrect data values; however, no significant changes occurred.

As already mentioned, the HIM is useful for stratifying homeless individuals into more precise categories than the traditional, dichotomous YES/NO variable. Homelessness is a transitory phenomenon; some people can be *more* homeless than others (O'Flaherty, 1996, p. 18). Thus the idea of the HIM is to attempt to explain the intensity of homelessness using demographic covariates. For example, a 50 year-old person who has been homeless for 6 months in his adult life would be much less intensely homeless than a 35 year-old who has spent 10 years of his adult life homeless. In the same fashion, individuals who regularly stay in a motel room but are street homeless for a few nights at the end of every month when their fixed benefits run out are much less intensely homeless than those who are homeless every night during the month.

Table 3.2 shows summary statistics for this important variable. On average, individuals in the greater homeless population are homeless for 23 percent of the nights since the beginning of their first homeless episode.

Figure 2.1 shows the estimated density of the homeless intensity measure. This is the estimate for the greater homeless population, constructed using the smaller, sampled population.

Table 2.6: Summary of HIM

Variable	**Mean**	**Min**	**Max**
Percent of nights unhoused since first episode began	0.23	0.05	1.0
N	157		

While the average of the HIM is 0.23, the mode is near 0.10 due to the long right tail. In other words, while it is common for individuals to have been homeless for about 10 percent of the nights since the beginning of their first homeless episode, some individuals (albeit a smaller proportion) have been homeless for nearly every night since their first homeless episode began.

In order to answer the question, "Given that an individual is homeless, what are the determinants of the intensity of homelessness?", the HIM is regressed on covariates that are clearly exogenous. For example, *age* and *gender* are used as covariates, but the endogenous *income* of the respondent is not. Analysis is restricted to adults; it is likely that children's homelessness is determined more by their parent's determinants than their own.

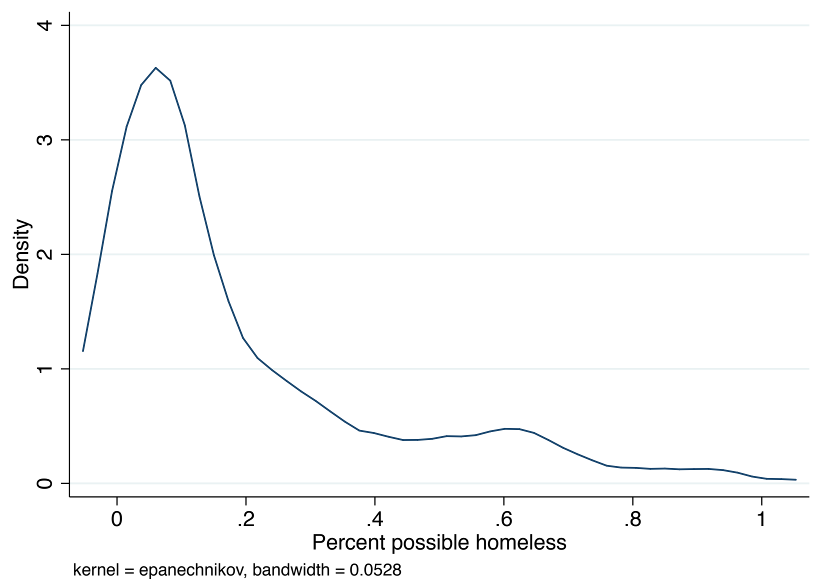


Figure 2.1: Density estimate for the homeless intensity measure (HIM)

2.5 Results

Table 2.7 reports the results of estimating the following regression:

$$HIM_i = \alpha + \mathbf{x}_i'\boldsymbol{\beta} + \varepsilon_i \tag{2.5}$$

where HIM_i is the homeless intensity measure for person i, and $\mathbf{x}_i$ is a vector that includes other covariates and demographic controls for person i.[9] Due to heteroscedasticity in the residuals, robust standard errors are used.

The coefficient of the male dummy is positive and significant. Even from the early ages of social science research, researchers have known that the homeless population is disproportionately male (see, for example, Anderson (1923)). According to these data, a male homeless individual can expect to spend 10 percentage points more of his time on the street than a female. Age, however, does not have an effect on the intensity of homelessness.

Having family in Costa Mesa (family that existed in CM before the episodes of homelessness began, which is therefore exogenous) exerts a negative and significant effect on the HIM. On average, individuals with family in CM spend 35 fewer percentage points of their nights on the street. This finding is expected, as an individual who has a family support system can rely on it to reduce homeless intensity.

Strong effects come from veteran status and combat status. In general, a homeless veteran is much more intensely homeless than a non-veteran homeless individual. Those in combat are a proper subgroup of the larger veteran group, but this indicator has an extremely negative

[9] With the homeless population in particular, it is likely that some of the data is reported with a degree of error, due to recall problems. However, in the data contained in the CM Assessment, the variables most likely to suffer from recall error are the measures of homeless intensity (such as, "How long have you been homeless?"). The variables used on the right hand side of equation (2.5) are much less likely to be recalled with error, as many times they are simple yes/no questions (e.g., "Have you ever served in the armed forces?") or oft-recalled facts ("How old are you?"). Thus, in the CM Assessment, measurement error will likely only occur in the HIM index. Fortunately, measurement error in the HIM (the outcome variable) does not bias $\hat{\boldsymbol{\beta}}$.

Table 2.7: Dep = % Possible Life Homeless

Variable	Coefficient	(Lin. Std. Err.)
Male	0.10***	(0.04)
Age	-0.01	(.01)
Family in CM	-0.35***	(0.10)
Child School CM	0.05	(0.04)
Foster	-0.10	(0.06)
Veteran	0.20***	(0.07)
Combat	-0.31***	(0.08)
Mental health	-0.10	(0.07)
No HS	*excluded*	
Some HS	-0.24*	(0.13)
GED	-0.30**	(0.15)
HS Graduate	-0.49***	(0.14)
Some College	-0.46***	(0.13)
College Grad	-0.42***	(0.14)
Post Grad	-0.50***	(0.15)
Freely Chosen	*excluded*	
Alcohol/Drug Abuse	-0.05	(0.13)
Job Loss/Econ Problems	-0.20	(0.10)
Injury/Health Problems	0.07	(0.12)
Evicted from House	-0.27***	(0.10)
Other	-0.27***	(0.11)
Relational Discord	-0.17*	(0.10)
Constant	0.94	(0.22)
Dep Var Mean	0.21	
N	114	
R^2	0.60	

$^{*}p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

effect. The combined effect, however, is not statistically significant, given that a test that the sum of the veteran and combat coefficients is zero (this is the effect for a combat veteran) fails to reject the null hypothesis of a zero value. It is likely that upon entering the armed forces, an individual's superiors observe his or her ability and make an assignment to a combat role or to a support role based upon the information gleaned during initial training. Other research has shown that combat soldiers are different (for example, more violent, both before and after joining the military) from those soldiers who do not see combat (MacManus et al., 2013). It therefore seems likely that those soldiers who have seen combat are different from non-combat soldiers in a way that is observed during basic military training and not observed in the CM Assessment. This unobserved characteristic is likely co-causing both their placement into a combat role, and their homeless intensity.

The educational attainment of an individual affects the intensity of homelessness in the direction that one would expect. For this variable, "No High School" is the omitted category. Predictably, Table 2.7 shows that the higher a homeless individual's educational attainment, the less intensely homeless he or she is (up until "Some College", before falling slightly). A high school graduate spends 49 fewer percentage points of his or her nights on the street compared to someone who has not graduated high school. Those who have completed some or all of a college degree show about the same effect (46 and 42 percentage points, respectively).

The self-reported cause of homelessness also appears to play a role. Specifically, those who are homeless because they have been evicted from their house are less intensely homeless. On average, people with this characteristic spend 27 fewer percentage points of their nights on the street. Those who are homeless due to relational discord also spend much less of their adult life homeless. This finding makes sense, as individuals in these categories probably weren't choosing homelessness the way some other individuals might; they were caught without a house as a result of a sudden life event.

An interesting component of these results is that two of the usually-proffered determinants of the drop into homelessness (which this paper does not attempt to answer directly) appear to have no effect on its intensity. Mental illness and being a graduate of the foster system (which was shown to have a positive effect in Piliavin et al. (1993)) both have insignificant coefficients in this model.

2.5.1 Check-In Center

In January 2012, in an effort to provide concrete solutions to the problem of homelessness in Costa Mesa, the Churches Consortium started a service for the homeless, called the Check-in Center. The Check-in Center is a safe place for individuals experiencing homelessness to "check-in" their belongings for the day in a secure location. The hope was that this option would free up the individuals from the detrimental effects of having to carry everything they own at all times, both in terms of damage to their bodies and the inability to seek employment. Table 2.8 shows selected usage patterns of the Check-In Center (hereinafter CI).

Table 2.8: Use of Check-In Center

	No	**Yes**	**Total**
Female	18	15	33
Male	65	29	94
Total	83	44	127
Veteran			
No	72	33	105
Yes	10	10	20
Total	82	43	125
Age			
18 to 30	14	1	15
31 to 50	31	14	45
51 to 61	26	21	47
62 and over	11	7	18
Total	82	43	125

Because use of the CI is not exogenous, being simultaneously determined along with the intensity of homelessness, it is necessary to instrument this variable with another excludable, relevant variable in order to appraise its effect on the intensity of homelessness. An indicator of familiarity with the CI was chosen as the instrument. This indicator is recorded as a one if the individual knew about the existence of the CI, and was recorded as a zero otherwise. This instrument is clearly relevant, as knowledge of the CI will affect the usage of it, and the first stage F-statistic is 12.06, making it reasonably strong (Staiger and Stock, 1997). It is also excludable; knowledge of the CI should not have any effect on the intensity of homelessness.[10]

Table 2.9 adds the CI variable to the model already presented in Table 2.7. It shows the two-stage weighted least squares (2SWLS) estimate of the effect of use of the CI on the intensity of homelessness. At -0.21, this is an extremely large reduction in the intensity of homelessness, given a dependent variable mean of 0.24. Therefore, it appears that the CI is very effective at reducing the percentage of nights spent homeless.[11] Furthermore, suppose there is doubt as to the appropriateness of the instrument. In this case, we would expect that use of the CI is likely positively correlated with intensity of homelessness; that is, as individuals become even more homeless, and begin to collect more of the trappings of homelessness (such as shopping carts, bed rolls, etc.), they are more likely to need and therefore use the CI. With failure to control for such simultaneity, the coefficient on the CI would then be positive. Therefore, the estimated negative coefficient suggests that the instrument is working as intended.

The rest of the coefficients are almost identical to those in Table 2.7, although the drop in

[10]The notion of excludability is eliminated if knowledge of the CI makes being homeless "better," or less bad, and this leads individuals to be homeless more often. However, time series snapshots of the quantity of homelessness before and after introduction of the CI do not seem to bear this out. Indeed, the CI was implemented after a similar (but much less exhaustive) census was conducted in 2010 and needs were assessed. The number of homeless remained fairly constant during this time. The Costa Mesa police, also, have no record of an increase in homelessness at the time of the introduction of the CI.

[11]If homeless individuals were found to use the CI only in the beginning of their homeless episodes and not in the later stages, this same result would be reported. Therefore, such usage patterns would be a threat to the validity of this result. However, I find no evidence of this pattern, nor have I found any in the associated literature.

the number of non-missing observations in this specification leads to larger standard errors, and thus less precision in the results. The most notable change is the coefficient on the gender dummy, which is now insignificant. No other variables, however, show large changes, and the reported R^2 values are identical to two decimal places.

Table 2.9: Dep = % Possible Life Homeless, 2SWLS

Variable	**Coefficient**	**(Lin. Std. Err.)**
CIuse	-0.21***	(0.08)
Male	0.05	(0.05)
Age	-0.01	(.01)
Family in CM	-0.15***	(0.05)
Child School CM	0.03	(0.05)
Foster	-0.09	(0.06)
Veteran	0.28***	(0.07)
Combat	-0.32***	(0.11)
Mental health	-0.07	(0.07)
No HS	*excluded*	
Some HS	-0.39***	(0.14)
GED	-0.52***	(0.17)
HS Graduate	-0.67***	(0.14)
Some College	-0.70***	(0.15)
College Grad	-0.67***	(0.17)
Post Grad	-0.50***	(0.15)
Freely Chosen	*excluded*	
Alcohol/Drug Abuse	0.05	(0.13)
Job Loss/Econ Problems	-0.17	(0.10)
Injury/Health Problems	0.03	(0.12)
Evicted from House	-0.29***	(0.10)
Other	-0.24**	(0.11)
Relational Discord	-0.17*	(0.10)
Constant	1.12	(0.22)
Dep Var Mean	0.24	
N	90	
R^2	0.60	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

2.6 Conclusions

In this paper, I utilize a novel data set collected in 2012 to investigate the determinants of homelessness intensity among those who are street homeless in Costa Mesa. By restricting the sample to only those who are already homeless and introducing the concept of a homelessness intensity measure, I exploit variation in individual characteristics and usage of a Check-In Center to predict the intensity of homelessness (defined as the ratio of actual homeless nights to potential homeless nights). By making use of a novel application of the weighted least squares estimator, I am able to estimate coefficients that are design-unbiased for the population values. By using an instrument, I find that the Check-In Center has a large negative effect on the intensity of homelessness. The self-reported cause of homelessness is a determining factor as well. I also find, as would be expected, that education lessens the intensity of homelessness.

Chapter 3

Individual Determinants of Homelessness: A Preference-Elicitation Approach

3.1 Introduction

This paper seeks to answer the following research question, "Given that an individual is homeless in Costa Mesa, California, what effect does this individual's discount rate and attitudes toward risk have on the intensity of homelessness experienced by that individual?" In what has been termed an "artefactual field experiment" (Harrison and List, 2004), individuals' risk preferences and discount rates (explained in more detail in Sections 3.4 and 3.5, respectively) are elicited in order to estimate their effect on individuals' intensities of homelessness (essentially the percentage of nights spent on the street, see Section 3.3). The paper finds that a decrease in patience and an increase in the level of risk tolerance (a more risky individual) are correlated with an increase in the level of homelessness. The direct effect of risk preferences on the intensity of homelessness is still significant even after controlling for observably risky behaviors such as smoking, drug use, jail time, etc. This means that an individual's risk preferences affect the level of homelessness in other ways than just through observably risky behaviors.

This paper is not seeking to answer the related question, "What causes homelessness?" Primarily, the current focus is due to the empirical limitations of the data collected (only unhoused, homeless individuals were surveyed), but it is also due to the belief of the author that there is no firm line delineating homeless and housed individuals. That is to say, among this population, there is a fair amount of churning about the line; individuals change housing status often.

Figure 3.1 demonstrates this idea empirically. This figure is the compilation of the 2013 homeless histories constructed for each homeless individual. We see that in 2013, of the homeless individuals interviewed, approximately 48 percent were not housed in January. This number dips to a low of about 43 percent in February, before climbing to a high of about 59 percent in October. This trend appears to be cyclical; the numbers continue to fall until December, and appear to be likely to achieve a low again in February, 2014 (a date that is outside the range of the 2013 homeless histories).

This trend, while doubtless influenced by a number of factors, appears to be strongly driven by the weather. October is a time of temperate weather in Costa Mesa, while the coldest nights are usually found in February (overnight lows occasionally dip below freezing). Anecdotally, the author has evidence of a number of homeless deaths during these cold nights. The problem is serious enough that the County of Orange, California (in which Costa Mesa is located), provides bus service to a nearby National Guard armory during the coldest nights so that homeless individuals from Costa Mesa can sleep more safely (Rojas, 2013).[1]

[1] The low numbers in February are not caused by the individuals' usage of the armory as a sleeping location; such individuals are still considered unhoused for the purposes of this paper.

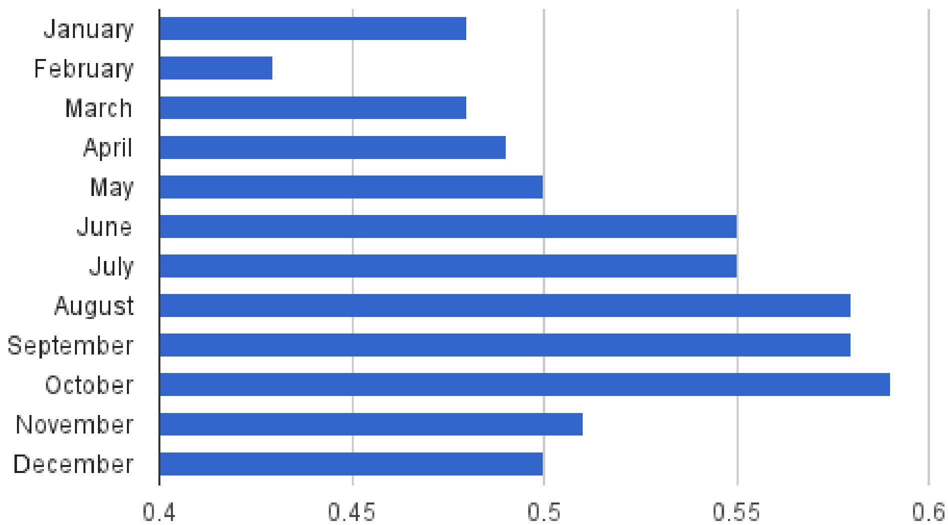


Figure 3.1: Unhoused population, as a percentage of total homeless population (by month, 2013)

Another way to understand Figure 3.1 is to imagine selecting a homeless individual at random. There is only a 43 percent chance that this individual will be found on the street (that is, unhoused) in February, while there is nearly a 60 percent chance this person will be on the street in October.

This paper is written from the perspective that the homeless population of interest is not merely the collection of individuals who are unhoused on any given night; rather, the population of interest is the complete data frame of all homeless individuals who have spent a night unhoused in a given time period (for this paper, the year 2013). This second number is much larger than the first, since this study suggests that the homeless in Costa Mesa spend, on average, 49.5 percent of the year without a roof (that is, unhoused). See Figure 3.2 for a hypothetical example. This paper would view the size of the full homeless population as three individuals (Bob, Barbara, and Joanne) regardless of whether the count is occurring on day one or day two.

Furthermore, seeing the homeless population from this vantage point means that the full effect of potential causal factors is not only the effect they have on the level of homelessness among those found on the street, but also the effect on those who have been/will be homeless yet are currently housed. Thus, the statistics presented in this paper originate from interviews conducted with homeless individuals who are currently unhoused, but the statistics have then been weighted in an attempt to represent this larger homeless population.[2]

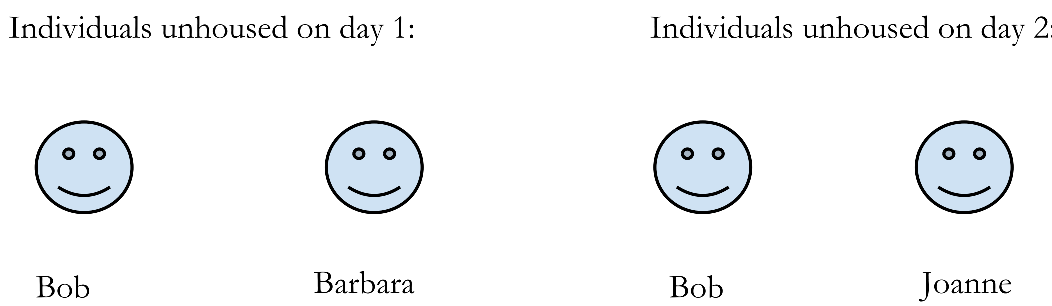


Figure 3.2: Hypothetical homeless population (n=3)

3.1.1 Comparison to author's previous work

This paper is, in many ways, an extension of Jarvis (2015)'s paper. That paper examines the individual-level determinants of homelessness from an earlier (2012) round of interviews of the homeless population in Costa Mesa. For the purposes of that paper (and in the current study), a homeless individual is someone who spent a night without a roof in the previous

[2]Further explanation and justification of weighting can be found in Section 3.2.3 on page 70.

year.[3] The earlier paper found that a church-administered intervention known as the Check-In Center had a large negative effect on the intensity of homelessness. The self-reported cause of homelessness was also a determining factor, and education and the presence of family nearby were found to lessen the intensity of homelessness. Approximately 180 people were interviewed; the paper also posited that the full population of homeless individuals in Costa Mesa was about 300 individuals.

In light of the aforementioned, this paper extends the first paper in three important ways. First, it conducts a similar full-census interview of Costa Mesa's homeless population 20 months after the first. This repetition is important in order to test the previous paper's hypothesis that approximately 180 people out of a full "population" of 300 were unhoused on any given night. By taking another snapshot of all the unhoused individuals 20 months later, the veracity of this previous claim can be tested. A comparison between the two censuses is found in Table 3.1.[4] As previously noted, both papers attempt to characterize the larger "homeless population" instead of characterizing only those who were unhoused at a given moment in time. If this characterization has been done correctly both times, the inferred populations' characteristics should be largely the same between the two studies while the composition of the unhoused populations themselves should be different in terms of individuals. This is because the unhoused individuals encountered during both studies are from the same larger homeless population, but due to the rapidly changing housing statuses of the members of this population (recall that on average, an individual in this population spends only 49.5 percent of nights unhoused), the *a priori* probability of being included in both interview groups is only about 24.5 percent.[5]

[3]There are varying definitions of homelessness. In both papers, someone who spent a night on the street, on a park bench, in a car, in a homeless shelter or "unhoused" in any other way would be considered homeless. Someone who moved into another's home ("doubling up") is not considered homeless.

[4]This table lists values for the "inferred" population; that is, the larger homeless population and not the unhoused found on any given night.

[5]Suppose a representative homeless individual named Xavier. Call event A the probability of X appearing in the first sample group. Since X is only unhoused 49.5 percent of the year, he would only be interviewed with a 0.495 probability. Thus, P(A) = 0.495. Call event B the probability of X appearing in the second sample group. By the same logic, P(B) = 0.495. Assume independence of the two events (being included

The 2014 study infers that the average age of the members of the total population is 51.5 years old, while the 2012 study infers an average age of 46.1. Recall that if the two studies are in fact making correct inferences about the total population, these numbers should be very nearly the same, with small differences attributable to sampling variation. According to the 2014 study, the members of the population have spent 4.2 years of their adult life homeless, and were 45.2 years old when they first became homeless, whereas the earlier study had estimated that the members of this game population had spent more time homeless (5.3 versus 4.2 years) and were younger (40.2 versus 45.2 years old) when they first became homeless. At 61 percent, the share of the population inferred by the 2014 study to be on their first homeless episode is larger than the share inferred by the earlier study, which is 49 percent. The average number of homeless episodes for each individual in the population is 2.3 for the 2014 study and 2.9 for the 2012 study. Also of note is the inferred population sizes in both studies. They are remarkably similar; even though different individuals (and different numbers of these different individuals) were interviewed, the 2012 study estimates 300 unique individuals and the 2014 study estimates 282 individuals in the larger homeless population.

While, as previously mentioned, the *a priori* probability of appearing in both unhoused samples is 0.245, the actual incidence of overlap of individuals between samples is 16 percent. This is somewhat lower than predicted, but not altogether unexpected for the following reasons. If either of the two surveys are incomplete censuses of the unhoused individuals, the probability of being included in either sample is less than the percentage of nights spent unhoused, and the intersection of these two events is therefore less than 0.245.[6] Furthermore, migration of individuals to other locations and/or transition to a permanently housed status (an exit from the homeless population) makes the actual overlap of individuals smaller than

in one interview round is unlikely to change the percentage of nights spent homeless). Thus, the $P(A \cap B) = 0.495 * 0.495 = 0.245$.

[6] While great care was taken to include all individuals who were unhoused during the interview period (as detailed in Section 3.2 on page 65), this undoubtedly occurred.

the *a priori* probabilities.

As for the second extension, the 2014 survey (which this paper analyzes) adds to the previous study by asking questions of interest that were not present in the earlier study. The 2014 study asks interviewees to self report objective indicators of brain injuries (such as a hospitalization, diagnosis, or the filing of a police report). A very high proportion (approximately 34 percent) of the homeless sample reported an objective indicator of a brain injury before becoming homeless the first time. This is much higher than the general population, which has an incidence of brain injury around 12 percent (Frost et al., 2013). However, it does not appear to be out of line with other studies of homeless individuals, such as Topolovec-Vranic et al. (2014), which found that 36 percent of men in a homeless shelter in Toronto screened positive for a brain injury that had occurred before their first homeless episode. As little empirical evidence currently exists examining the linkages between brain injury and homelessness, this is likely an important area of future research.

Another new type of question in the 2014 survey designated events as occurring either before or after the first episode of homelessness. As seen in Table 3.1, 72 percent of the individuals in the population had been incarcerated prior to their first homeless episode, and 83 percent had attended a drug or alcohol treatment program.

Finally, and most substantially, this paper adds to the previous paper by examining the effect of the two variables of interest on the intensity of homelessness. This paper seeks to estimate the effect of an individual's risk aversion and discount rate on the intensity of homelessness.

Table 3.1: Summary and comparison of censuses

	Inferred Pop 2012	Inferred Pop 2014
Age	46.1	51.5
Age at first episode	40.2	45.2
Total years homless	5.3	4.2
Times homeless	2.9	2.3
First time homeless	0.49	0.61
Male	0.72	0.75
Foster	0.12	0.06
Veteran	0.13	0.20
% Adult life homeless	0.23	0.16
% 2013 homeless	*n/a*	0.50
Brain Injury†	*n/a*	0.34
Ever-incarcerated†	*n/a*	0.72
Drug/alcohol treatment†	*n/a*	0.83
Present in 2012 census	*n/a*	0.16
Number of interviews	180	162
Inferred Pop Size	300	282

† measured before (or at time of) first homeless episode

3.1.2 Literature review

In addition to the Jarvis (2015) paper mentioned earlier, a small but informative literature exists which examines the determinants of homelessness at the individual level. Rolston et al. (2013) find that the presence of a homelessness prevention program in New York City (known as Homebase) reduced the number of shelter entries by over 60 percent. However, Goodman et al. (2014), studying the same program, concluded that the average length of a homeless family's shelter stay was unaffected by participation in Homebase.

Braga and Corno (2011) interviewed 62 percent of the street homeless individuals in Milan and found that the population is overwhelmingly male and middle aged (which this study also finds). Piliavin et al. (1993) found, in what is perhaps the most similar study to the one presented in this paper, that among a non-random sample of 330 homeless individuals in Minneapolis, the duration of homelessness is positively influenced by less consistent work histories and participation in the foster system, and is not influenced by symptoms of alcoholism.

3.2 Methods and sample creation

The first step of this study, and perhaps the most important/logistically difficult one, was to construct and then interview a representative sample of the population of interest (those who have spent at least one night unhoused in Costa Mesa in 2013). Careful study of the homeless population is difficult due to a dearth of good quality data. Not only are homeless individuals hard to find and track, but many of these potential interviewees feel that authority figures in their lives have taken advantage of them, making them extremely reluctant to sit down and talk with such people. Diligent attention was paid to these factors in the survey design process.

One additional advantage of having the universe of unhoused individuals (instead of the at-the-time unhoused individuals) as the population of interest is that data is collected for all potentially homeless individuals, even ones who answer “No” when asked if they are homeless. This is useful because these same individuals, when prompted through a homeless history of the previous year, often report being unhoused, even going so far as to report being currently unhoused (perhaps due to embarrassment when initially asked about their housing status). With this survey method, the data for all individuals is collected. An objective determination is then made after the fact regarding their inclusion into the population of interest, based on the existence of an unhoused night in the previous year.

From experience gained during the earlier 2012 round of surveys, it was decided that instead of actively searching for and identifying homeless individuals, interviewers (trained undergraduate students from the University of California, Irvine’s Undergraduate Research

Opportunities Program) would be stationed at all service provider locations around the city for a period of time to wait for the homeless individuals to come to them.[7] Given a long enough time period, it is likely that most, if not all of the homeless individuals visited at least one of the service provider locations.

After communicating with and receiving permission from the service providers in the city, the undergraduates were stationed at each one in order to conduct interviews. The service locations, for their part, posted signs and explained the survey process in the weeks prior to the survey. For example, if a certain church hosted a soup kitchen from 5-8pm on a Tuesday, the soup kitchen would announce the upcoming survey for one or two weeks prior to the appearance of an interview team, who would attend during the normal soup kitchen hours the following week. Every effort was made to make the undergraduates appear to be "part of" the agency that they were stationed at. Since most service locations survive by using a constantly rotating volunteer staff, incorporating the undergraduates into the volunteer corps was not difficult. This was done in order to build trust among the homeless interviewees: they already trusted the service provider, and from the perspective of the interviewees, the interviewer was a member/volunteer of the service location. Building trust was crucial for this population; not only were many of the questions sensitive, but also many homeless individuals avoid trouble by being "invisible".

The general sampling methodology consisted of waiting at a service provider location until a homeless individual was about ready to leave. An interviewer would then approach this homeless individual and ask for his or her participation in the survey. More often than not, the homeless individual already knew about the survey, and would consent. In all cases, and especially when the individual did not know about the interview and/or was reluctant to participate, the interviewee would be informed that all participants had the opportunity to earn between four and 16 dollars by consenting. All individuals at the service

[7] These locations included churches, soup kitchens, social service agencies, etc.

provider locations had the opportunity to participate, regardless of housing status (but their housing status and homeless histories were updated to reflect the differences). Currently housed individuals (current as of the interview date) who had been previously homeless were included because they were still members of the population of interest; individuals who had never been homeless and were therefore not in the population of interest were still allowed to participate and earn remuneration so there would be no incentive to falsely report housing status to earn a higher payoff.

The teams arrived at a different service location every day to interview individuals. After the interview was complete, the interviewees received their payment and a blue wristband which they were asked to wear to indicate that they had been interviewed already. One member of the interview team was assigned to take names (and compare them to the list of individuals interviewed thus far) as the individuals arrived in order to dissuade potential interviewees from being interviewed more than once in order to receive a higher payoff.[8] The interview teams continued to attend a different service location every day until all potential interviewees had been interviewed, at which point the study was considered to be complete. After eight days of interviewing, over 220 interviews had been conducted. After removing from the data set individuals who had never been homeless, children under 18, and homeless individuals who did not pass a Costa Mesa "residency" requirement,[9] the trimmed data set included data for 162 individuals. By sharing this list with leaders of the service organizations and other social workers, and by recording the number of refusals, it appears that this survey methodology successfully interviewed between 90 and 95 percent of the individuals unhoused at the time.

One additional challenge with this population was the possible lack of mental acuity in an

[8]Since wristbands can be taken off, this additional step was taken to ensure that no one was interviewed more than once.

[9]This "residency" requirement was met by having attended (self or children) school in Costa Mesa, having ever rented or owned a home in Costa Mesa, having worked or lived in Costa Mesa immediately prior to the first homeless episode, and/or having family who lived in Costa Mesa at the moment the individual became homeless the first time.

unknown subset of the interviewees. Therefore, in this study, the trained undergraduates entered the data into the survey forms, instead of the respondents themselves. This effort reduced the presence of incorrect data in the data set. The undergraduates were instructed to read the question exactly as written the first time through, and provide basic guidance if needed (such as the definition of a word, or converting an answer from days to weeks if the prompt specified such a time period). The undergraduates were instructed not to discard answers they thought might be incorrect, as this could potentially introduce bias into the data.

The interview instrument itself was converted into a Google form, which was then accessed by the interview team using their cellphones and/or tablets. As the interviewers entered data, a spreadsheet was automatically created that merged the details of each interviewee into one secure document.

Unfortunately, even the meticulous sampling method described above is subject to (at least) three main issues:

- Potential bias due to migration after losing housing
- Potential bias due to selection of convenient subjects
- Need for adjustment of sampled population to match target population

This study attempts to ameliorate each of these issues, as explained in the following sections. The result is that the estimates presented in this paper can be treated as if they come from a representative, random sample of the homeless population.

3.2.1 Migration bias

Since the goal of this sampling methodology was to create a sample representative of all the people who are unhoused in Costa Mesa in a given time period, if many Costa Mesa homeless individuals leave the city after becoming homeless, then bias would be introduced into the analysis. It is likely that the migrants' location decisions would be correlated with an unobservable, such as mental aptitude, that also influences homeless intensity. The result would be bias in the estimated coefficients of the covariates determining the intensity of homelessness.

It is therefore important to note the selection of Costa Mesa as the survey site. Costa Mesa can best be thought of as a "net receptor" of homeless individuals. Costa Mesa's mix of climate and services make it unlikely that a homeless individual would migrate away once becoming homeless. Indeed, one of Costa Mesa's pressing concerns (as outlined in the bulletin City of Costa Mesa, 2012) is the prevention of non-Costa Mesa homeless taking up residence in Costa Mesa while simultaneously providing services to only those homeless who are Costa Mesa residents. Furthermore, nearly 80 percent of the interviewees identified Costa Mesa as their primary residence before becoming homeless. While this does not, in fact, prove the contrapositive (that most Costa Mesa people who become homeless stay in Costa Mesa), it lets the reader understand the nature of migration flows among the homeless. Due to the aforementioned reasoning, by interviewing all the homeless in Costa Mesa and removing from them the non-Costa Mesa residents, we can arrive at a sample that is (largely) free of migration bias.

3.2.2 Selection bias

Due to the difficulty of sampling this population, convenience sampling is very often used. For example, much of the individual-level homeless data reported by the Department of

Housing and Urban Development is reported to them by care agencies who sample those whom they can, with the count then scaled up to match the size of the whole population. This is known as snowball convenience sampling, in which the interviewee leads the team of researchers to another potential interviewee, who leads them to another, and so on. The discriminating reader can understand why this would lead to selection bias, as it is likely that the selection of interviewees in this scheme is correlated with an unobservable characteristic that might also be correlated with the outcome variable, intensity of homelessness. This bias, while at times unavoidable in the world of homelessness research, is overcome in this study due to the sampling scheme: the (almost) simultaneous surveying of (almost) all homeless individuals in Costa Mesa.

By surveying every individual on the street at a given time, the selection biases inherent in the other methods are eliminated. This approach was only possible because Costa Mesa is a relatively small city ($\approx$16 mi^2) with a relatively small street homeless population. Furthermore, if a homeless individual doesn't want to be found, it is unlikely that he or she will be found. But by waiting at the common service locations (soup kitchens, etc.), the researchers in this study were able to interview the "recalcitrant" types when they came to replenish supplies. Thus, due to the unique characteristics of Costa Mesa, it is likely that the survey sample did not suffer from selection bias.

3.2.3 Weighting to create a representative sample

As mentioned earlier, the population of interest for this study is the universe of all individuals who spent one night unhoused in 2013, but the sampled population is those who were currently unhoused at the time of the interview procedure. Therefore, it is necessary to make an adjustment to the sampled population to more accurately reflect the target population.[10]

[10]Please see Jarvis (2015) for a more technical discussion of the weighting procedure.

The ideal scenario would be to look at the target population and, with identical and independent probabilities of selection, select a sample of individuals. This approach is the foundation for a representative sample. However, in the sample selection procedure described above, an individual in the target population will only be chosen if the interviewers are interviewing at the same time this individual is unhoused. In other words, individuals who were unhoused every night in 2013 had a 100 percent chance of being included in the sample, whereas individuals who were only unhoused for one night had an 8 in 365 sample inclusion probability.[11]

The solution to this problem is to adjust data for all of the individuals so that they have the same sample inclusion probability (Pfeffermann, 1996). If the individual with an 8/365 sample inclusion probability is weighted by the reciprocal of his or her sample inclusion probability, then this person's sample inclusion probability becomes $(8/365) * (365/8) = 1$, which is equal to the sample inclusion probability of the other individual. By the same logic, all individuals must be weighted by the inverse of their sample selection probability.

Weighting in this fashion also allows "inclusion" of those individuals who are in the larger homeless population but are currently housed. The sampling methodology described above only interviews those individuals who are currently unhoused and does not include those in the larger homeless population who are currently housed. Weighting these individuals (who were not unhoused for very many nights in 2013 but who did happen to be unhoused at the time of the interview period) by a larger amount allows them to stand in for the individuals who were not able to be interviewed because they were housed. In other words, by making the aforementioned, one-night-homeless individual count as 365/8 people, he or she "counts" for all the other individuals who were unhoused at some point in 2013 but were not unhoused at the time of the interview process.

[11] Since the interview process took eight consecutive days and this supposed individual was unhoused for only one day, he or she had eight chances to be included in the sample (the eight days of interviews) and 365-8=257 chances to not be included in 2013.

It is especially important to weight the individuals according to their sample selection probability if we suspect that some outcomes of interest differ according to that probability, as is the case here. For example, suppose we want to know the average age of everyone in the full homeless population (the target population, i.e., those who have spent at least one night in 2013 homeless), and we estimate this value using the average age of everyone in the sampled population (only those who were unhoused during the interview process). Suppose further that all young homeless people are unhoused only a few nights of the year, while older homeless individuals are unhoused every night of the year. Then, for any day that interviews take place, more older individuals will be interviewed than young ones. That is, the share of older individuals in the sampled population is much higher than their actual share of the target population, because a large quantity of the younger homeless are housed that day. Thus, the value estimated for average age will be biased upward because the older homeless individuals are over-represented in the sampled population.

The need for weighting becomes even more pronounced in a policy context. Suppose that the government crafts a policy to reduce nights spent on the street and that this policy has a large, beneficial effect on these younger homeless while having no effect on the older homeless. Further suppose that the magnitude of this effect is estimated from data collected using an unweighted sample of homeless individuals. Then the policy will appear to be useless, when in fact it is very useful; it is useful to the few young, homeless individuals who were sampled as well as to the large number of young, homeless individuals who were not sampled because the were not unhoused at the time of the interview. In order to properly estimate the actual effects of this policy, it is necessary to weight the younger individuals more than the older ones to make up for the many young homeless individuals who were not sampled.

Define, generally, w_i as the inverse of the sample inclusion probability for person i. Conveniently, it can be shown that $E_D(\sum_{i \in S} w_i) = N$, where $E_D(.)$ indicates the design expectation over all the possible samples that could be generated using the sampling scheme, S is the set of all sampled individuals, and N is the total number of homeless individuals in the larger homeless population. In essence, taking the expectation over the design means that the statistic would asymptotically converge to this number if the sample were drawn from the population an infinite number of times, and the statistic were calculated each time (see Pfeffermann (1996); Carrington et al. (2000) for further explanation).

Furthermore, the estimator

$$\frac{1}{N}\sum_{i\in S} x_i w_i \tag{3.1}$$

is therefore the best estimate for any population parameter x, and should be used instead of $\bar{x}$. However, since N is not known,

$$\frac{1}{\sum_{i\in S} w_i}\sum_{i\in S} x_i w_i \tag{3.2}$$

is the best feasible estimator of any population parameter x and is used in computing the statistics in Table 3.1.

In similar fashion, Wooldridge (1999) shows that the presence of varying inclusion probabilities means that the usual OLS estimates from a regression carried out on the sample will be inconsistent because they assume equal probability of selection.[12] However, Wooldridge finds that weighting the criterion function (in the OLS case, the sum of squares) by the sample inclusion probability restores consistency. Weighting the criterion function by w_i is

[12] Solon et al. (2013) clarify, explaining that when the sampling probabilities are functions of the independent variables, OLS estimates are still unbiased. They explain that it is only when the inclusion probabilities are correlated with the outcome variable that OLS is biased. This latter case is the situation in this paper.

acheived by simply rewriting the population regression so that:

$$\sqrt{\boldsymbol{W}}y = \sqrt{\boldsymbol{W}}\boldsymbol{X}\boldsymbol{\beta} + \sqrt{\boldsymbol{W}}\boldsymbol{\varepsilon}, \tag{3.3}$$

where $\boldsymbol{W}$ is an nxn matrix with w_i on the ith diagonal. From here, it is easy to see that the OLS estimator must be replaced by a weighted least squares estimator, $\hat{\boldsymbol{B}}_{\boldsymbol{wls}} = (\boldsymbol{X'WX})^{-1}(\boldsymbol{X'WY})$.

However, the standard errors in this case are not the standard errors generated by the usual weighted least squares estimator that is used more frequently to deal with heteroskedasticity, as here we weight by the inverse of the root of the inclusion probability, instead of the inverse of the root of the variance. Kott (1991) provides a linearization estimator that is nearly unbiased for the variance of $\hat{\boldsymbol{B}}_{\boldsymbol{wls}}$, and this correction is used in this paper.

For the aforementioned reasons, the weighted least squares estimator shown above is the appropriate choice to estimate effects on the whole population and will be used in this paper. The percentage of nights spent unhoused in 2013 is used as the sample inclusion probability.[13] [14]

Therefore, by the arguments presented in the previous sections, the sample can be treated as a random sample of the homeless population of Costa Mesa, and as such the estimates are valid for population inference.

[13]This is true with one caveat: The probabilities were truncated to fall between five and 100 percent, inclusive. Probabilities that fell outside this interval were replaced with the nearest endpoint.

[14]This is also the same measure that is used in the formulation of the HIM, explained in the next section.

3.3 Homelessness intensity measure

In this section, one of this research program's contributions to the literature on homelessness is described: the concept of a homeless intensity measure (HIM). This paper utilizes the following simple measure: The percentage of nights in 2013 spent unhoused. As discussed in the previous section, the percentage of nights spent unhoused is also the individual's sample selection probability. Thus, the HIM ends up being the inverse of the w_i weight mentioned earlier. As a result, regression observations are weighted by the root of the inverse of the HIM.

Specifying the HIM in such a way is useful for stratifying homeless individuals into more precise categories than the traditional, dichotomous YES/NO variable. Homelessness is a transitory phenomenon; some people can be *more* homeless than others (O'Flaherty, 1996, p. 18). Thus the idea of the HIM is to attempt to explain the intensity of homelessness using demographic covariates. For example, a 50 year-old person who has been homeless for 6 months in his adult life would be much less intensely homeless than a 35 year-old who had spent 10 years of his adult life homeless. In the same fashion, individuals who regularly stay in a motel room but are street homeless for a few nights at the end of every month when their fixed benefits run out are much less intensely homeless than those who are homeless every night during the month.

Table 3.2 shows a summary of the HIM used in this paper. It shows that, as previously discussed, the average homeless individual was unhoused for 49 percent of 2013.

Table 3.2: Summary of HIM

Variable	**Mean**	**Min**	**Max**
Percent of nights in 2013	0.49	0.08	1.0
N	157		

Due to the possibility of recall error, a detailed homeless history was only constructed for the year 2013 (instead of an individual's entire homeless past). In order to construct the most accurate homeless history possible for each individual, the interviewees were asked a series of questions about their housing status on certain dates (such as, "Were you homeless on July 4th?", "Were you homeless on your birthday?", "Were you homeless on Thanksgiving?", etc.) in order to provide anchor dates around which they could more easily determine the dates of their homeless episode.[15]

In order to answer the research question posed earlier, "Given that an individual is homeless in Costa Mesa, what effect does an individuals' risk aversion and discount rate have on the intensity of homelessness?", the HIM is regressed on the risk aversion and discount rate measures, along with other covariates that are clearly exogenous as controls. For example, *age* and *gender* are used as covariates, but the endogenous *income* of the respondent is not. The analysis is also restricted to adults over 18; it is likely that children's homelessness is determined more by their parents' determinants than their own.

3.4 Risk preferences

A commonly proffered driver of homelessness is the propensity of an individual to undertake risky behaviors. At times it is supposed that some individuals are more homeless than others because of the tendency to gamble away money, or abuse drugs, or commit a crime which leads to a prison sentence, or become employed in a risky occupation with a high morbidity rate (such as construction or military service). In essence, each of these suppositions appear to be attempting to explain that an individual's level of homelessness is influenced by a higher-than-average affinity for risk. In the economics context, we explain differences in risk

[15] Though every effort was taken to minimize it in the construction of the homeless histories, recall error does not pose a serious threat to the analysis in this paper. Since the HIM is the left-hand-side variable in the regressions (y_i), non-systematic error in this variable is subsumed into the regression equation's own error term. Non-systematic error in the right hand side variables, however, would bias the coefficients on those variables toward zero.

attitudes by differences in the concavity of the underlying utility function, as in Figure 3.3. These suppositions form the *ex-ante* hypothesis of this paper, and provide the motivation for examining risk preferences.

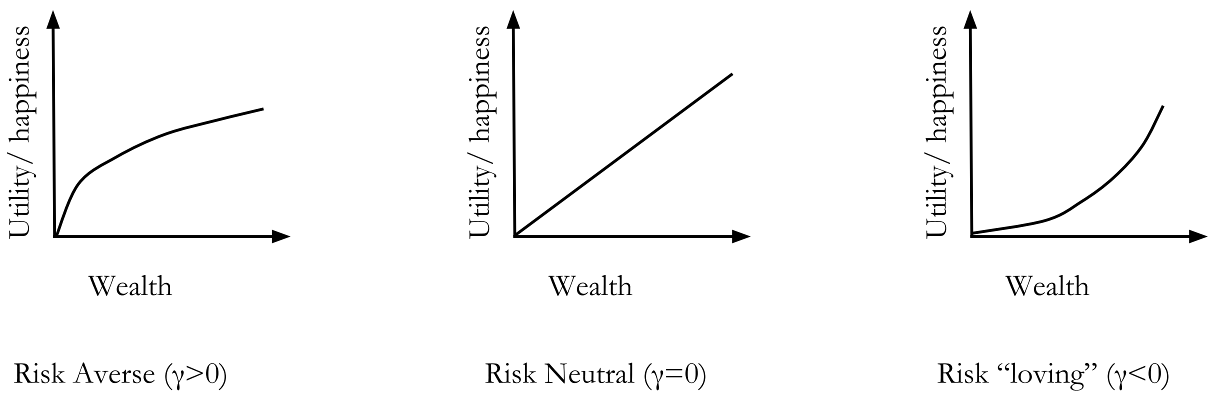


Figure 3.3: Underlying utility functions which generate associated risk tendencies

This paper tries to elicit the underlying attitudes towards risk (and the corresponding utility functions) that the individuals possess, in order to see if they have a direct effect on an individual's HIM. This is important in a policy context in the following fashion. Suppose that a person's attitudes toward risk directly influence the level of homelessness (which is a finding of this paper) in addition to any secondary, indirect effects that come through risky behaviors. In this scenario, the common practice of attempting to control an individual's risky behavior may not lead to a large reduction in the level of homelessness (if it reduces it at all), as the homelessness is also being driven directly by the individual's risk preferences.[16] However, if the opposite is true, then it might be the case that behavior modification can lead to improvements in the level of homelessness. In either situation, it is important to answer this question empirically if possible.

The regression analysis in Section 3.6 seems to suggest that an individual's risk attitudes

[16]Practices such as these are very popular in the faith-based, homeless-aid community. For example, many shelters require their attendees to abide by strict codes prohibiting alcohol, smoking, gambling, etc. in order to qualify for benefits.

directly influence the level of homelessness to a great extent. It also appears that the risk attitudes indirectly influence the level of homelessness by first influencing observably risky behavior, which then influences the level of homelessness.

3.4.1 Mathematical foundation of risk preferences

In keeping with current literature (Andersen et al., 2008; Dave et al., 2010; Holt and Laury, 2002), a constant-relative-risk-aversion utility function is used, of the form:

$$u(c) = \begin{cases} \frac{1}{1-\gamma} c^{1-\gamma} & : \gamma \neq 1 \\ \ln(c) & : \gamma = 1 \end{cases} \tag{3.4}$$

This utility function is useful because the marginal utility is always positive and diminishing:

$$u'(c) = c^{-\gamma} > 0, \quad u''(c) = -\gamma c^{-(\gamma+1)} < 0. \tag{3.5}$$

The level of γ here is important. Using the standard measure of relative risk aversion, $-cu''(c)/u'(c)$, the function in equation (3.4) implies risk aversion equal to γ. Furthermore, as can be seen in equation (3.5), the greater is γ, the more negative the second derivative, and the greater the concavity of the utility function. In other words, individuals with higher levels of γ are more risk averse. In this way, γ is commonly referred to as the coefficient of relative risk aversion. Thus, it is important to say something about the level of γ implied by each individual's choice of gamble.

This paper bounds each individual's level of γ in the following way. Knowing that each individual will select the gamble that maximizes expected utility, and knowing that the most risk averse (highest γ) individuals will select the least risky gamble, the question is asked, "At which level of gamma would an individual be indifferent between the least risky gamble and the subsequent (slightly more risky) gamble?" The level of γ which solves:

$$5 = \int_{4.25}^{6.25} \frac{1}{1-\gamma} c^{1-\gamma} \frac{1}{6.25 - 4.25} \, dc \tag{3.6}$$

is the boundary value of γ for the individuals who chose the least risky gamble (see Table 3.3 for payoffs). In equation (3.6), the left hand side is the expected utility of the sure bet (the least risky gamble available), and the right hand side is the expected utility of the next gamble, calculated using $E(L) = \int l f(l) \, dl$. As the gambles are all over continuous uniform distributions, l is a payoff that can take any real value between the two endpoints of the lottery played and $f(l)$ is the continuous uniform probability density function between the same two endpoints. For individuals with a value of γ greater than the one which solves equation (3.6), the sure bet is strictly preferred to all other gambles. The same reasoning is used on each successive pair of gambles, giving buckets within which each individual's level of γ must lie.

Table 3.3: Risk Games (subject chooses which to play)

Choice (unif. dist.)	Low Endpoint	High Endpoint	Expected Return	Standard Deviation	Implied CRRA Range
A	5.00	5.00	5.00	0.00	$\gamma \geq 8.11$
B	4.25	6.25	5.25	0.58	$8.11 > \gamma \geq 2.67$
C	3.50	7.50	5.50	1.15	$2.67 > \gamma \geq 1.60$
D	2.75	8.75	5.75	1.73	$1.60 > \gamma \geq 0.00$
E	2.25	9.25	5.75	2.02	$0.00 > \gamma \geq -1.07$
F	1.00	10.00	5.50	2.60	$-1.07 > \gamma$

This approach is followed for two reasons. The first is to describe what fraction of homeless individuals choose risk-"loving" ($\gamma < 0$) gambles, and the second is to compare the levels of γ to those found in subjects participating in other experiments. This paper differs slightly from the previously mentioned papers in that the risk gambles played are not 50/50 gambles, but rather continuous uniform distributions with progressively higher variances. Thus, determination of γ took place using the method outlined in equation (3.6) on the previous page.[17] The exact details of each gamble are outlined in Table 3.3 on the preceding page.

The implied range of the coefficient of relative risk aversion, γ, is shown for each gamble. Table 3.3 shows that an individual with a γ of 5 (a relatively high level of risk aversion) would have preferred to play gamble B. Thus it can be understood that the value of γ for any individual that chose to play gamble B lies between 2.67 and 8.11, since every individual will choose to play the gamble that maximizes his or her utility. An individual whose value of γ is 0 is risk neutral; this same individual would be indifferent between playing gambles D and E, but strictly prefers those two gambles to every other. Finally, someone with a value of γ more negative than -1.07 can be considered risk "loving"; that is, he or she prefers uncertainty. This individual would choose gamble F. This gamble is notable in that the expected return is *less* than that of gamble E, but the variance is greater. Individuals who chose this gamble value the ability to win ten dollars (the high end point) more than the associated loss in expected return that occurs from selecting this gamble.

The validity of this elicitation scheme hinges on the ability of interviewees to understand the gamble choices. For this exercise, interviewees were given a visual description of each gamble on a sheet of letter paper. On this paper, six number-lines were drawn, one for each gamble, and the lines were labelled with the appropriate letter (A, B, C, D, E, F) to indicate which gamble they represented. On each number line, the high endpoint, low endpoint, and expected value of the gamble was listed. The undergraduate interviewers explained to the interviewees that they could choose any gamble that they wanted, and they would get a payoff partly depending on which gamble they chose and partly depending on chance. The interviewees were helped to understand that the payoff could be anywhere (in \$.01 increments) between the high and low endpoints with equal probability, and that the average

[17] Continuous distributions were used, instead of 50/50 gambles, at the request of the Institutional Review Board, who asked the author to devise a way to pay each individual an equal amount of money, regardless of which gamble they had chosen.

payoff was just the number in the middle, between the two endpoints. The interviewers were instructed to explain the game to the individuals and answer their questions. Once each interviewee understood the task, he or she made a decision, and the interviewer recorded the letter of their choice digitally into the spreadsheet and physically on a card. The interviewee then took the card to the undergraduate in front of the room, where he or she was given an envelope with the letter of their selected gamble written upon it. They were assured that the envelope contained some value between the endpoints of the selected gamble, that is, the randomized winnings.

Although all subjects received the same payout (five dollars) due to the request of the Institutional Review Board, measures were taken to foster the illusion of heterogenous payoffs so that the game would accurately elicit risk preferences.[18] This included placing various combinations of dollar bills and coins (each combination, however, summed to five dollars) into the envelopes, asking individuals to wait until they had left the testing area to open the envelopes, and requesting that they not share the quantity of their winnings.[19]

Table 3.4 on the next page shows the proportion of homeless individuals that selected each gamble. The results in this table (with the exception of the college comparison column) are weighted using the technique described earlier. Over half (59 percent) of the homeless individuals selected one of the two riskiest gambles (E or F). This indicates that these individuals are of the risk "loving" sort, that is, given two gambles with an equal expected payoff, they will always choose the riskier one. This group of people pertain to the right panel in Figure 3.3. Males and females both were very likely to choose either of the two most risky gambles, with females slightly more likely to choose gamble F. On the other end of the scale (and pertaining to the left panel in Figure 3.3), 13 percent of the homeless population chose the least risky gamble. This gamble paid out five dollars with certainty, and males

[18] If, at any time, the individuals realized that the payoffs were identical, they would no longer care which gamble they chose, and their chosen gamble would no longer be an accurate reflection of their risk preferences.

[19] It turns out that this last point was self-enforced. None of the homeless individuals were willing to divulge the quantity of their winnings for fear of becoming a target of theft.

were slightly more likely to take it than the females. In general, however, male and female homeless individuals chose similarly, with clear preferences, on average, for the more risky gamble.

Table 3.4: Preference Elicitation

	Female Homeless	Male Homeless	All Homeless	College Comparison
Risk Game				
A (no risk)	0.11	0.14	0.13	0.43
B	0.11	0.04	0.06	0.00
C	0.16	0.12	0.13	0.17
D	0.06	0.10	0.09	0.09
E	0.16	0.44	0.37	0.04
F (most risky)	0.39	0.16	0.22	0.26
Discount Rate				
Now	0.77	0.71	0.72	0.09
Later	0.23	0.29	0.28	0.91
Now | Yes attend	0.56	0.72	0.69	*n/a*
Later | Yes attend	0.43	0.28	0.31	*n/a*
N	34	82	116	26

For purposes of comparison, and to make sure the answers from the homeless population were believable, the same risk-aversion-elicitation method was utilized with a randomly selected group of 26 college students. Much work has been done eliciting risk preferences from college students; it is therefore easy to compare the results from the college students in this study to results from studies conducted with very similar populations. In this study, the median college student chose gamble C, which corresponds to a value of γ somewhere between 1.60 and 2.67. Dave et al. (2010) and Holt and Laury (2002) estimate that the average individual in the population has a γ of approximately two. Because the results from the college group comparison are reasonable and in line with what previous research indicates, it is likely that the risk aversion elicitation technique described above accurately describes an individual's preferences towards risk.

The difference in risk attitudes between the homeless population and the college comparison group is remarkable. Whereas 59 percent of the homeless population demonstrated risk-“loving” behavior, only 30 percent of the college comparison showed this tendency. Furthermore, only 13 percent of the homeless population chose the sure bet; 43 percent of the college comparison did so. A naïve comparison of these two groups seems to indicate a significant difference in the risk attitudes of the two groups and therefore a potential link between risk aversion and homelessness.

Table 3.5 on the following page further investigates this idea, by comparing the risk level with the level of homelessness. Given the *ex-ante* hypothesis presented earlier, we would expect that the more risky an individual, the higher his or her HIM would be. The table shows a trend that moderately agrees with this hypothesis. The individuals who chose the least risky game on average spent 30 percent of 2013 unhoused, which rises to a maximum of 70 percent for those who chose risk game D, and then falls slightly to 45 percent for those who chose the most risky game. This very apparent trend provides a theoretical justification for including an individual's risk preferences in a regression explaining homeless intensity. The data in this table can also be useful when interpreting the magnitude of the effects in Table 3.6 on page 93, as the HIM is the dependent variable in the regression reported in that table.

We would also expect that properly elicited risk preferences should be highly correlated with observably risky behaviors. Table 3.5 provides support for this view, as well. Fifty-two percent of the individuals who chose the lottery with no risk had attended a drug or alcohol treatment before their first homeless episode. This number follows a fairly well defined trend upward, peaking at 100 percent for those who chose lottery D, before falling slightly to 95 percent for those who chose the most risky lottery. In general, the more risky the lottery that was chosen, the higher the probability of having attended a drug or alcohol rehabilitation center before the first homeless episode.

The ideal experiment would be to measure an individual's risk preferences on January 1, measure this individual's HIM at the conclusion of that year, and then regress the HIM on the risk preferences. This paper, however, measures individuals' risk preferences and then calculates their HIM for the preceding year.[20] If the intensity of homelessness in the year 2013 influences the individuals' risk preferences, there exists a threat of endogeneity. Thus, identification in this paper is dependent upon the assumption that risk preferences for homeless individuals are static in the short run. This is only a mild assumption; risk preferences are allowed to change when an individual undergoes an extreme shock such as his or her descent into homelessness for the first time. However, once this change has occurred, given that the individual is already homeless, this paper assumes that his or her risk preferences don't change as their homeless intensity changes. This is similar to what Ohmura et al. (2006) found, that is, a short-term stability in preferences.

Table 3.5: Results by Risk level (weighted)

Risk game	Percent 2013 Unhoused (HIM)	Probability of having attended drug/alcohol treatment
RISK A (no risk)	0.30	0.52
RISK B	0.26	0.80
RISK C	0.32	0.82
RISK D	0.70	1.00
RISK E	0.66	0.86
RISK F (most risky)	0.45	0.95

3.5 Delay discounting

The other variable of interest examined in this paper is an individual's discount rate (also known as the level of future orientation, and/or the intertemporal rate of substitution). In other words, this is the measure of the extent that an individual dismisses, or "discounts" a future payoff in exchange for an immediate payoff. An individual with a high discount rate can be thought of as impatient: a future payoff of 100 dollars means much, much less to

[20]Eliciting a preference and comparing it to a previous behavior is common practice in the laboratory literature, see Bickel et al. (1999); Dixon et al. (2003) for examples.

this person than an immediate payoff of 100 dollars; an individual with a low discount rate values both payoffs more equally.

Both academic research and anecdotal evidence seem to indicate that a person's discount rate can influence certain behaviors. Anecdotally, it is sometimes supposed that a homeless individual's situation might be influenced by the lack of investment in human capital (such as a degree or other training) and/or the lack of investment in financial assets that could be utilized when the individual undergoes a negative shock (such as the loss of a job, illness, or the deterioration of a relationship). When this shock hits, individuals with accumulated assets can perhaps utilize these assets to ward off homelessness, while individuals who did not accumulate the assets fall more quickly into an intensely homeless episode. One of the main explanations for why some people accumulate assets while others do not, all else equal, is a difference in discount rates. Individuals with a low discount rate prefer the future payoff that their accumulated assets will provide more than they prefer present consumption, while individuals with a high discount rate prefer present consumption. Therefore, this paper investigates the effect that an individual's discount rate has on their level of homelessness. The hypothesized mechanism (and the *ex-ante* hypothesis of this paper) is that a decrease in the discount rate reduces the intensity of homelessness as described above.

Academic research appears to agree that discount rates can influence behavior (but little or no research has been conducted examining the effect on homelessness in particular). Chabris et al. (2008) find that laboratory-measured discount rates predict behaviors at least as well as standard predictor variables such as demographic characteristics. Steinberg (2007) (speaking about adolescents but with a developmental neuropsychology framework that can be applied to adults) explains a multitude of behaviors using variations in discount rates. Steinberg et al. (2009) elaborate, showing experimentally that for subjects over the age of 16, a preference for future payoffs (versus current ones) is driven primarily by variation in discount rates and not just a difference in impulse control. Johnson and Bickel (2002) and

Madden et al. (2003) provide a justification for laboratory-style elicitation of discount rates, as they show that individuals' reactions to real-world payoffs (such as a college degree) are similar to their reaction to experimental payoffs (such as those used in this paper). This is important, as it suggests that the discount rates elicited in this paper will be an accurate reflection of the individual's true future orientation. Finally, Ohmura et al. (2006) show a remarkable stability of an individual's short-term discount rate, which can help the results of this study be robust to endogeneity of the following sort. If discount rates evolve over time, and/or an individual's homeless status changes an individual's discount rate (the converse of what is claimed in this paper), then any demonstrated effect of the discount rate on the level of homelessness will not capture causality. However, Ohmura et al. (2006) conclude that an individual's short-term discount rate is stable (certainly for the time horizon, two to three months, containing the duration of the majority of homeless episodes). Given this, the results in this paper, which find evidence that cross-sectional variation in the discount rate affects the level of homelessness, are likely robust to this concern of endogeneity.

Having empirical evidence of the magnitude of the effect of the discount rate on the level of homelessness is important for policy selection. Many homeless interventions implicitly assume that homeless individuals have levels of future orientation that are similar to those of the policy makers themselves. For example, a homeless intervention policy that offers job training to homeless individuals so that they can obtain a job and then lift themselves out of homelessness may be reasonable to a policy maker and palatable to his or her constituents. However, this type of program requires significant up-front costs from the participants in terms of their time and effort, but it does not pay off until much later. For an individual who discounts the future greatly, the future benefit of employment is much less than the current costs of attendance, and the individual does not enroll in the program. Many homeless intervention programs have this property; very few, with the exception of the new wave of "housing first" interventions (see Buchanan et al. (2009) for an example) provide an immediate benefit. If, as this study suggests, homeless individuals on average discount the future more, and this propensity to discount the future is associated with an increase in the level of homelessness, then homeless prevention programs that both require a significant investment of resources now and don't offer a payoff until much later will be ineffective. That is, from the perspective of a large proportion of the homeless population's members, participation in the program will make them worse off than non-participation.

3.5.1 Discount rate elicitation

This study elicited the discount rate of homeless individuals by offering them a choice between two compensation schemes. The individuals were offered the choice between four dollars immediately at the end of the interview, or eight dollars the following day. Because the interviews occurred every day at various locations, individuals who wanted the later payment were able to collect it at the interview session that was to occur the next day. Individuals who value the future equally with the present (these are the low-discount-rate individuals) were more likely to choose the delayed, and greater, payment than the other individuals. Using the same logic, high-discount-rate individuals were more likely to opt for the "now" payment. Their choice was recorded by the interviewer, who recorded it digitally and on a card, and the card was taken to another interviewer at the front of the room whose job it was to disburse money. If they had selected the delayed payment, they were instructed to come the following day to the interview location. Cards with directions were also provided (but were often unneeded, as the homeless individuals had a pretty good idea of the locations of service providers).

Table 3.4 on page 82 shows the summary of the discount-rate elicitation. The majority of respondents (72 percent) selected the "now" payment instead of the later payment. This rate of uptake was fairly even for men and women, with women (at 77 percent) slightly more likely than men (71 percent) to select the "now" payment. When comparing this group with the college comparison group, the difference is large but not altogether unexpected. Only nine percent of college students selected the now payment. Given that college students are currently engaged in a very costly process whose payoff is many years in the future, we would expect college students to have a very low discount rate, or they would forego college. The agreement of the results with this hypothesis serves to confirm the validity of the elicitation methodology.

However, there still exist two additional threats to the validity of the elicitation methodology previously described. The following paragraphs detail these threats, along with the measures taken to eliminate them and tests to check the effectiveness of these measures.

The first threat can be described as "future-state uncertainty". If the individual did not know with 100 percent the state of the world that would be revealed the following day, they would be more likely to select the "now" payment even though they were of the patient type. This type of uncertainty might also have been generated by an individual who was unsure about being able to find the location the following day to receive payment. Whatever the cause of the uncertainty, it could be controlled for by offering the individuals the opportunity to receive the later payment at a location they were sure to be present at the following day. If the individual was assured of being able to receive the future payment, individuals who selected the "now" payment would only do so due to differences in discount rate. This paper asked the individuals, "Do you regularly attend ________ ?" where the blank was filled in with the name of the subsequent day's interview location.[21] Regularly attending the interview location served as the proxy for the case in which the individual was guaranteed to be able to receive the future payment wherever they happened to be the following day.

The last two rows of Table 3.4 on page 82 show the breakdown of "now" versus "later" choices conditional on having answered "yes" to the question above. If future-state uncertainty

[21]This question was asked very early in the survey, long before the preference-elicitation tasks had begun, in order to ensure that the interviewees did not understand why it was being asked. This helped guarantee truthful elicitations.

was influencing the decision of people during the discount-rate-elicitation task, it would influence these individuals to choose the "now" payment over the "later" payment due to the uncertainty of being able to collect the "later payment". Thus, when restricting analysis to the sample that assured the interviewers that they would be at the interview location the following day (this is a group that does not suffer from future-state uncertainty), the proportion of people selecting the "now" payment should drop considerably. For the male subgroup, the proportion of individuals selecting the "now" payment actually increases, which is evidence against the idea that future-state uncertainty is driving peoples' decisions. For the female subgroup, the number does drop significantly. This is perhaps evidence that the female subgroup does suffer from this uncertainty. Certainly, it might be the case that female homeless individuals are less in charge of their daily routines, especially if they are homeless with another person of the opposite sex who may determine these routines.[22] Therefore, in the regression analysis in the following section, the effect of the discount rate on the level of homelessness is analyzed separately for each gender group, owing to the fact that the females' responses to the discount rate elicitation task might have been affected by future-state uncertainty.

The second threat to the validity of the discount-rate-elicitation task is a lack of trust on the part of the interviewees about the realization of the future payment. If the interviewees did not believe that the interviewer was trustworthy, they would pick the "now" payment over the future payment due to uncertainty regarding the possibility of obtaining the future payment. Thus, the homeless population would appear to discount the future more than they actually do.

This threat is neutralized through careful selection of the interview location and setting. By partnering with the various service providers, the interviewers were able to appear to the homeless individuals as if they were volunteers of the service provider. This allowed

[22]Pairings (informal marriages and/or families) are common among the homeless, partly due to a desire for protection (Smith, 2008). It is possible that females give up some level of autonomy for this protection.

the homeless individuals to trust the interviewers from the beginning, since the homeless individuals already trusted the service provider as an institution. Furthermore, the homeless community is not very large. If the interviewers withheld money or engaged in activities otherwise deemed inappropriate, word would have spread quickly, encouraging the homeless individuals to distrust the interviewers. In the absence of such a phenomenon, the homeless individuals would feel more comfortable selecting the future payment. Indeed, some potential interviewees declined interview requests in the first few days in order to watch the proceedings more carefully, only to accept an interview request at a later date/location once they had deemed that the interviewers were trustworthy.

This study utilizes a simple check to determine if these measures were sufficient to imbue the potential interviewees with enough confidence to select the "later" payment if appropriate. Since the interviews happened sequentially, one per day at different locations, but many of the same individuals were cognizant of the proceedings each day, as the interview process carried on, the pool of potential interviewees learned more about the interviewers and were able to increase their level of trust. If the earlier individuals were "trust-constrained" and picked the "now" payment due to this constraint and not due to their discount rate, then the later individuals, not constrained by the lack of trust, would have selected the "now" payment with a lower frequency than the earliest individuals.[23] Since this is not the case (there is no discernible trend in the data as the days go on), we can conclude that the careful integration of the interviewers into the service location was sufficient to satisfy the trust needs of this part of the elicitation process.

[23] This is true with two caveats: First, later individuals cannot be more trusting than earlier individuals, and second, there can be no association between the individuals' discount rates and the time (early in the process versus late) that they were interviewed.

3.6 Results

Table 3.6 on page 93 shows the results of the main regression. Three models are shown. The first uses all controls and is run using the full sample.[24] The second model is restricted to men only, as women may have had some trust issues with the discount rate elicitation process (discussed in Section 3.5). Finally, the third model removes observably risky behavior, to analyze the direct versus indirect effects of risk aversion (as discussed in Section 3.4).

As expected, risk preferences do affect the HIM. In this regression, a dummy for risk "loving" behavior is utilized as the measure of risk.[25] The individuals who chose one of the two most risky gambles, the ones that indicate a negative γ (and therefore an upward sloping utility function of total wealth as shown in Figure 3.3 on page 77), are coded with a one. The others are coded with a zero. We see that in every model, having chosen one of the riskiest gambles is associated with a great increase in the HIM: 13 percentage points in the first model, 29 and 40 percentage points, respectively for models two and three. Comparing the model in column two with the model in column three, we see that omitting the observably risky behaviors greatly increases the effect that the risk-"loving" dummy appears to have. This is likely a bias created by omitting these variables. In other words, the risk preferences of the individuals affect their levels of homelessness directly and indirectly (indirectly, by first affecting observably risky behaviors, which then affect the level of homelessness).

Also as expected, the individuals with a lower discount rate, that is, those who had selected the "later" payment, were much less homeless. Having a low discount rate decreased homelessness by 16 percentage points for the whole sample, and about 30 percentage points for

[24] In addition to the covariates shown, the following variables were present in the regression: current age and education. The following were used, with values as of their first episode of homelessness: dummies for heart disease, diabetes, asthma, cancer, kidney disease, liver disease, treatment for mental health issues, and alcohol abuse.

[25] This effect shown is robust to many different ways of including risk preferences, even including all six of the risk games as a categorical variable. This latter technique, however, is much more noisy and less significant.

men only. The coefficient on the dummy variable indicating a brain injury prior to the first homeless episode is negative and insignificant. This coefficient was expected to be positive and significant since this population appears to have a much higher incidence of head injury than the general population. Males, on average, spent 27 percentage points more nights homeless than their female counterparts. This is in line with, but greater than, the 10 percentage point increase Jarvis (2015) found in the same population in 2012.

Not only is foster care a risk factor for adult homelessness, but many foster children have homeless parents as well (Zlotnick et al., 1998), making this a strong predictor of homeless intensity. Table 3.6 shows that for the entire homeless population, having been in foster care leads to an increase of 23 percentage points of the number of nights spent homeless. Individuals who are older at the time of their first homeless episode are also more likely to be more intensely homeless. Having been incarcerated or having participated in a drug or alcohol treatment before becoming homeless the first time increases the level of homelessness by 19 percentage points (incarceration) and 27 percentage points (drug/alcohol treatment). Finally, those who were homeless for the first time on 2013 were much less likely to be intensely homeless. This finding likely indicates that it is easier to end a homeless episode if it is the first homeless episode; subsequent episodes become harder to end, and the likelihood of chronic homelessness increases.

3.7 Conclusion

This paper studied two posited determinants of the intensity of homelessness: an individual's level of risk aversion and his or her rate of discounting of future payoffs. Individuals were interviewed during an 8 days period in 2014 to provide data to study these effects. In order to create a sample that was representative of the whole homeless population, instead of only being representative of the individuals unhoused at the time of the interviews, a weighting

Table 3.6: Regression Results

	% Nights in 2013	% Nights in 2013	% Nights 2013
Risk “loving”	0.13*	0.29***	0.40***
	(1.76)	(3.11)	(4.38)
“Patient”	-0.16**	-0.27***	-0.32***
	(1.90)	(-3.29)	(-3.60)
Brain injury†	-0.12	-0.06	-0.03
	(-1.54)	(-0.84)	(-0.45)
Male	0.24**	—	—
	(2.59)		
Foster	0.23**	0.31*	0.12
	(2.10)	(1.94)	(0.62)
Age†	0.02***	0.02***	0.02***
	(3.24)	(2.97)	(2.89)
Ever-incarcerated†	0.19**	0.27**	—
	(2.27)	(2.52)	
Drug/alcohol treatment†	0.31***	0.10	—
	(2.84)	(0.80)	
First time homeless	-0.24***	-0.54***	-0.60***
	(-3.09)	(-5.25)	(-5.53)
Genders included?	Both	Male only	Male only
Observably risky behaviors?	included	included	not included

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

† measured before (or at time of) first homeless episode

strategy was employed.

The findings supported the original hypotheses, which were the following: First, an increase in risk-"loving" preferences leads to an increase in the intensity of homelessness. Second, an increase in the rate at which an individual discounts the future (a less patient individual) also leads to an increase in the intensity of homelessness. These variables of interest were selected because they are often used to try to explain homelessness anecdotally. Therefore, these findings are of special interest to policy makers when deciding between different homeless interventions, many of which implicitly assume certain characteristics about the risk and discount preferences of homeless individuals.

Bibliography

Andersen, Steffen, Glenn W. Harrison, Morten I. Lau, and Elisabet Rutstrom (2008), "Eliciting risk and time preferences." *Econometrica*, 76, 583–618.

Anderson, Nels (1923), *The Hobo: The Sociology of the Homeless Man*. University Of Chicago Press, Chicago, IL.

Attanasio, Orazio, Emla Fitzsimons, Ana Gomez, Martha Isabel Gutiérrez, Costas Meghir, and Alice Mesnard (2006), "Child education and work choices in the presence of a conditional cash transfer programme in rural colombia." *IFS Working Papers*, 13.

Bank, World (2006), "Project appraisal document on a proposed loan in the amount of us$60.0 million to the republic of ecuador."

Basu, Kaushik and Pham Hoang Van (1998), "The economics of child labor." *The American Economic Review*, 88, 412–427.

Bickel, W, Odum A, and Madden G (1999), "Impulsivity and cigarette smoking: delay discounting in current. never, and ex-smokers." *Psychopharmacology*, 146, 447–454.

Braga, Michela and Lucia Corno (2011), "Being homeless: Evidence from italy." *Giornale degli Economisti e Annali di Economia*, 70 (Anno 124), pp. 33–73.

Buchanan, D., R. Kee, L. Sadowski, and D. Garcia (2009), "The health impact of supportive housing for hiv-positive homeless patients: A randomized controlled trial." *American Journal of Public Health*, 99.

Carrington, William J, John L Eltinge, and Kristin McCue (2000), "An Economist's Primer on Survey Samples." Working Papers 00-15, Center for Economic Studies, U.S. Census Bureau.

Chabris, Christopher F, David Laibson, Carrie L Morris, Jonathan P Schuldt, and Dmitry Taubinsky (2008), "Individual laboratory-measured discount rates predict field behavior." *Journal of Risk and Uncertainty*, 37, 237–269.

County, Orange (2013), "Homeless count and survey report." Technical report, Focus Strategies, Sacramento, CA.

Dave, Chetan, Catherine C. Eckel, Cathleen A. Johnson, and Christian Rojas (2010), "Eliciting risk preferences: When is simple better?" *Journal of Risk and Uncertainty*, 41, 219–243.

Diallo, Yacouba, Frank Hagemann, Alex Etienne, Yonca Gurbuzer, and Farhad Mehran (2008), "Global child labour developments: Measuring trends from 2004 to 2008." Technical Report 1, International Labour Office.

Dixon, M, J Marley, and E Jacobs (2003), "Delay discounting by pathological gamblers." *Journal of Applied Behavior Analysis*, 36, 449–458.

Edmonds, Eric (2006), "Child labor and schooling responses to anticipated income in south africa." *Journal of Development Economics*, 81, 386–414.

Edmonds, Eric and Norbert Schady (2008), "Poverty alleviation and child labor." *Policy Research Working Paper Series*.

Frost, R.B., T.J. Farrer, M. Primosch, and D.W. Hedges (2013), "Prevalence of traumatic brain injury in the general adult population: A meta-analysis." *Neuroepidemiology*, 40, 154–159.

Goodman, Sarena, Peter Messeri, and Brendan O'Flaherty (2014), "How effective homelessness prevention impacts the length of shelter spells." *Journal of Housing Economics*, 23, 55 – 62.

Harrison, Glenn W. and J. A. List (2004), "Field experiments." *Journal of Economic Literature*, 42, 1013–1059.

Holt, C. and S. Laury (2002), "Risk aversion and incentive effects." *American Economic Review*, 92, 1644–1655.

Honig, Marjorie and Randall K. Filer (1993), "Causes of intercity variation in homelessness." *The American Economic Review*, 83, pp. 248–255.

HUD (2009), "The 2009 annual homeless assessment report to congress." Technical report, U.S. Department of Housing and Urban Development.

Jarvis, Justin (2015), "Individual determinants of homelessness: A descriptive approach." *Journal of Housing Economics, forthcoming*.

Jeliazkov, Ivan, Jennifer Graves, and Mark Kutzbach (2008), "Fitting and comparison of models for multivariate ordinal outcomes." *Advances in Econometrics: Bayesian Econometrics*, 23, 115–156.

Jencks, Christopher (1995), *The Homeless*. Harvard University Press.

Johnson, M. and W. Bickel (2002), "Within-subject comparison of real and hypothetical money rewards in delay discounting." *Journal of the Experimental Analysis of Behavior*, 77, 129–146.

Kott, Phillip S. (1991), "A model-based look at linear regression with survey data." *The American Statistician*, 45, pp. 107–112.

MacManus, Deirdre, Kimberlie Dean, Margaret Jones, Roberto J Rona, Neil Greenberg, Lisa Hull, Tom Fahy, Simon Wessely, and Nicola T Fear (2013), "Violent offending by uk military personnel deployed to iraq and afghanistan: a data linkage cohort study." *The Lancet*, 381, 907 – 917.

Madden, G., A. Begotka, B. Raiff, and L. Kastern (2003), "Delay discounting of real and hypothetical rewards." *Experimental and Clinical Psychopharmacology*, 11, 139–145.

of Costa Mesa, City (2012), "Assessment and implementation strategy." URL `http://www.costamesaca.gov/modules/showdocument.aspx?documentid=10028`.

O'Flaherty, Brendan (1996), *Making Room*. Harvard University Press, Cambridge, MA.

Ohmura, Y., T. Takahashi, N. Kitamura, and P. Wehr (2006), "Three-month stability of delay and probability discounting measures." *Experimental and Clinical Psychopharmacology*, 14, 318–328.

Pfeffermann, Danny (1996), "The use of sampling weights for survey data analysis." *Statistical Methods in Medical Research*, 5, 239–261.

Piliavin, Irving, Michael Sosin, Alex H. Westerfelt, and Ross L. Matsueda (1993), "The duration of homeless careers: An exploratory study." *Social Service Review*, 67, pp. 576–598.

Project, National Radio (2009), "How we survive." URL `http://www.radioproject.org/2009/02/how-we-survive-the-deepening\-homeless-crisis/`.

Quigley, John M. and Steven Raphael (2001), "The economics of homelessness: The evidence from north america." *European Journal of Housing Policy*, 1, 323–336.

Ravallion, Martin, Shaohua Chen, and Prem Sangraula (2009), "Dollar a day." *The World Bank Economic Review*, 23, 163–184.

Ravallion, Martin and Q Wodon (2002), "Does child labour displace schooling? evidence on behavioural responses to an enrollment subsidy." *The Economic Journal*, 110, C158–C175.

Rojas, Rick (2013), "Costa mesa confronts cold head on." URL `http://articles.latimes.com/2013/jan/19/local/la-me-homeless-cold-20130120`.

Rolston, Howard, Judy Geyer, and Gretchen Locke (2013), "Final report: evaluation of the homebase community prevention program." Technical report, ABT Associates, Bethesda, MD.

Schady, Norbert and Maria Caridad Araujo (2006), "Cash transfers, conditions, school enrollment, and child work : evidence from a randomized experiment in ecuador." *Policy Research Working Paper Series*.

Smith, Hilary (2008), "Searching for kinship: The creation of street families among homeless youth." *American Behavioral Scientist*, 51, 756–771.

Solon, Gary, Steven J. Haider, and Jeffrey Wooldridge (2013), "What are we weighting for?" Working Paper 18859, National Bureau of Economic Research.

Staiger, Douglas and James H. Stock (1997), "Instrumental variables regression with weak instruments." *Econometrica*, 65, 557–586.

Steinberg, L. (2007), "Risk taking in adolescence: New perspectives from brain and behavioral science." *Current Directions in Psychological Science*, 16, 55–59.

Steinberg, Laurence, Sandra Graham, Lia O'Brien, Jennifer Woolard, Elizabeth Cauffman, and Marie Banich (2009), "Age differences in future orientation and delay discounting." *Child Development*, 80, 28–44.

Topolovec-Vranic, Jane, Naomi Ennis, Mackenzie Howatt, Donna Ouchterlony, Alicja Michalak, Cheryl Masanic, Angela Colantonio, Stephen W. Hwang, Pia Kontos, Vicky Stergiopoulos, and Michael D. Cusimano (2014), "Traumatic brain injury among men in an urban homeless shelter: observational study of rates and mechanisms of injury." *Canadian Medical Association Open Access Journal*, 2, E69–E76.

Wooldridge, Jeffrey M. (1999), "Asymptotic properties of weighted m-estimators for variable probability samples." *Econometrica*, 67, 1385–1406.

Zlotnick, C, D Kronstadt, and L Klee (1998), "Foster care children and family homelessness." *American Journal of Public Health*, 88, 1368–1370.