|  |  |
| --- | --- |
|  | **The Effects of Using Minor Drugs During Adolescence: A Logistic Regression Approach** |
|  |  |
|  | Malcolm Bonner  Machine Learning |

**I. Introduction**

Drug use is a major health issue in our society today. The number of drug-related deaths due to overdose has steadily increased over that last decade. According to the Center for Disease Control the rate of deaths per 100,000 has risen from below 4 in 2000 to 15 per 100,000 in 2014. This an alarming trend and juxtaposed to data from the National Institute of Health proclaiming that over $700 billion is lost annually due to health care cost, crime and lost productivity. If policy indicatives are to be enacted to mitigate this issue, then there must be adequate research on the causes of this behavior to direct policy. There has been much research and debate around what role the gateway effect plays in an individual’s transition into drug use. In this paper I will explore this effect by using a logistic regression approach to discern whether or not an adolescent’s choice to take drugs has any effect on the probability that they will consume drugs as an adult. I hypothesize there will be a correlation between adolescent drug use and drug use as an adult. I also, hypothesize that there will be a difference based on the age in which the adolescent is introduced to drugs, as well as when the adolescent was born. This is due to changes in the legality of certain drugs such as alcohol and changes in drug culture.

**II. Data**

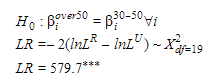
In order to complete this analysis, I will the use the National Survey on Drug Use and Health. The NSDUH is a national survey done annually on randomly selected individuals ages 12 and older. “Data from the NSDUH provide national and state-level estimates on the use of tobacco products, alcohol, illicit drugs (including non-medical use of prescription drugs) and mental health in the United States” (NSDUH).

**III. Methodology**

I will explore the probability that an adolescent using a minor drug is likely to consume a hard drug as an adult. In order to do this, drugs must be partitioned into minor and hard categories. For the purposes of this paper, alcohol, tobacco and marijuana will be classified as minor drugs. The reasoning behind this is these are the three most popular drugs among adolescents and most addiction literature views these three drugs as the gateway drugs that possibly lead to harder drug use. I will define hard drugs as cocaine or crack, heroin and methamphetamines. I omit prescription drug, because they were not prevalent until more recently in the sample. These will be partitions serves as the categories for dependent variable of the logistic regression. A person will fall into one of three categories they will be a hard drug user, minor drug user or a non-drug user.

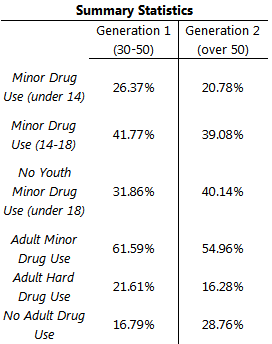
Before beginning the regression, the data must be trimmed and partitioned into generations. First, all people that are younger than 30 years of age are dropped from the sample. The reasoning behind this is to not have any bias in the data based on an individual not having been an adult for a long enough period of time. Most individuals will consume hard drugs during their twenties and I wish to have every individual in the sample to have lived through their twenties so the most accurate estimates can be obtained.

Next, I conduct a structural break test on the two generation to make sure doing the partition makes statistical sense. The rationale for choosing the generation partition at age 50 is this would distinguish between the changing of the alcohol laws and the increased social acceptance of drugs in the late sixties and early seventies. I conduct a likelihood ratio test with the null hypothesis being that the two generations are the same and thus only a pooled model should be considered.



As shown above the likelihood ratio test is above the threshold from the Chi-Squared distribution table so, the partition at age 50 seems to have significance.

After completing all the trimming and partitioning the following sample is obtained:



The summary statistics are broken up into two parts, adolescence and adulthood and presented in percentage terms. For example, in Generation 1, 26.37% of the sample used a minor drug before the age of 14, 41.77% of the sample used a minor drug between the ages of 14-18, lastly 31.86% consumed no drugs under the age of 18. Some trends are initially apparent. Drug use as a whole seems to be more prevalent in the younger generation than in the older generation.

**Model**

Logistic regression is a modelling technique similar to Ordinary Least Squares regression, but instead of predicting the value of a continuous dependent variable, logistic regression allows for the prediction of probabilities of an observation being in a category. In its essence, it is a classification technique. It makes use of the logistic function can take any value from negative infinity to positive infinite and return a value bounded between 0 and 1, which is useful for interpreting output as a probability.

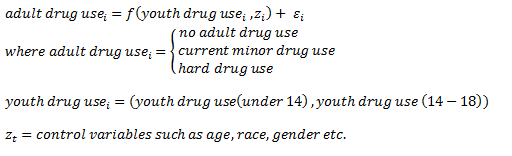
*t* can then be expressedas a linear combination of the explanatory or independent variables.

So, the logistic function is composed as follows:

Then the inverse logistic function or logit function is used in order to serve as the link between the regular regression and the probability space, such that the following expression is obtained:

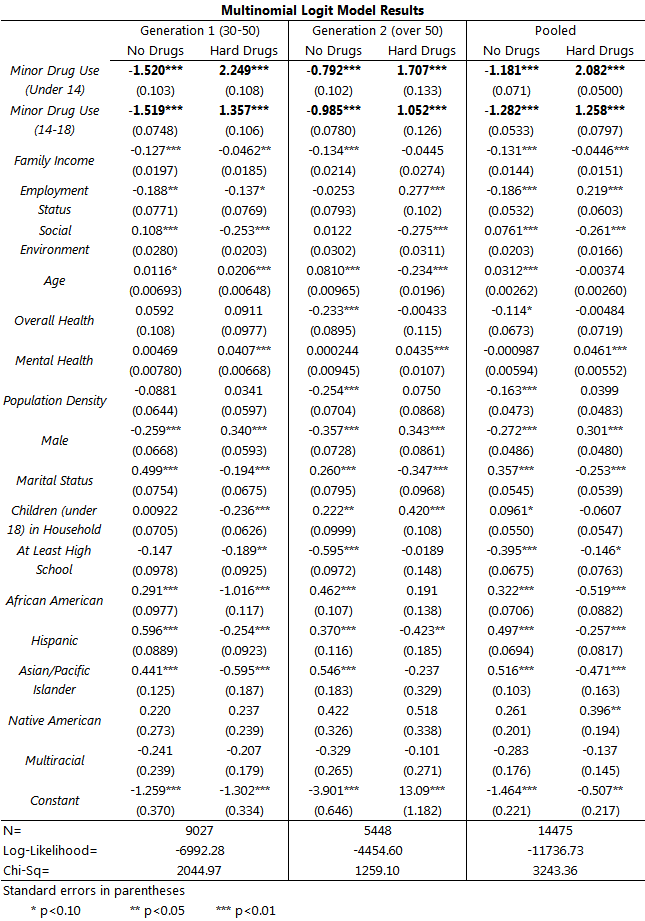
“The odds of the dependent variable equaling a case (given some linear combination x of the predictors) is equivalent to the exponential function of the linear regression expression” (Freedman).

The model for this paper will be a multinomial logit, rather than the normal binomial logit, because there are three categories for the dependent variable: no adult drug use, adult minor drug use and adult hard drug use. In order to use the multinomial method one of the categories must be chosen as a base to which the other categories will be compared. I will choose adult minor drug use, because it would signify that an individual’s drug use behavior did not change from adolescence to adulthood. This will make the interpretation of the results more intuitive than any of the other choices. The main explanatory variable will be the individual’s use of minor drugs as an adolescent and this will also consist of two indicator variables representing whether the individual consumed minor drugs before age 14 and those who consumed minor drugs from the age of 14-18. In addition to the primary explanatory variables, I also use many control variables consisting of race, age, gender, family income etc.



**IV. Results**

The following are the results from performing the above regression analysis:



The result of the regression seems to confirm the hypothesis that there is a correlation between minor drug use as an adolescent and hard drug use as an adult. The model finds in Generation 1 (30-50) using minor drugs under the age of 14 causes the odds that the individual does no drug as an adult to decrease by 1.52 relative to continuing minor drugs and the odds that the individual does hard drugs as an adult increase by 2.249 units relative to the continuing of minor drugs. These coefficients are both found to be significant at the 1% level, meaning this is a strong correlation. It is important to note that statements can only be made about correlation and not causation. These individuals may have inherent or environmental factors that lead them to drug use, but I attempt to control for many of these external factors with the inclusion of the numerous control variables.

Across the generations, the data show that there has been an increase in the effects from the consuming drugs as an adolescent on an individual’s drug use behavior as an adult. In Generation 2(over 50), the odds of moving to hard drugs after having done mild drugs under the age of 14 increases by 1.707 units relative to the continued use of minor drugs. This is in contrast to the increase of 2.249 units in Generation 2. The effect is also seen in the decrease in the probability that an individual will not consume drugs as an adult from Generation 2, where the odds are -0.792, to Generation 1, where the odds are -1.52 units. This indicates that drug behavior is not only escalating to a greater degree, but it is also stickier meaning, once people start doing any drug they are more like to continue than the previous generation. Similar results can be found for individuals who took minor drugs between the ages of 14-18, although the effects on that population of escalating to hard drug as an adult where not quite as strong as though whom consumed minor drugs before the age 14. This difference in effect makes sense given common knowledge about addiction escalation and tolerance build up.

**V. Conclusion**

In conclusion, this paper began with two hypotheses, mainly, adult hard drug use is structurally different across generations and that there is a correlation between adolescent drug use and adult drug use. The analysis in this paper using a multinomial logistic regression method shows there is a structural difference in drug use behavior between adults over age 50 and adults between the ages of 30 and 50. I speculate that this due to the changes in governmental alcohol policies that increased the age of alcohol consumption and the increase in social acceptance of drug use in the 1970’s. The data also show that there is a correlation between using drugs as an adolescent and an individual’s probability of using hard drugs as an adult. There is an increased probability that those individuals whom use minor drugs under the age of 18 will either continue to use minor drugs or escalate to some form of hard drugs. The implications of these results for policy makers are to focus on delaying, as much as possible, the consumption of minor drug by a child.

**Works Cited**

1. Ellickson, P. & Morton, S. (1999). Identifying adolescents at risk for hard drug use: racial/ethnic variations. *Journal of Adolescent Health,* 25: 382-395

2. Kenkel. D., Mathios, A.D. & Pacula, R. (2001). Economics of youth drug use, addiction, and gateway effects. *Addiction,* 96: 151-164

3. Kandel, D. & Yamaguchi, K. (1993). From beer to crack: developmental patterns of drug involvement. *American Journal of Public Health June*. Vol. 83, No. 6, pp. 851-855.

4. National Survey on Drug Use and Health, 2011

5. Pacula, R. (1997). Economic Modeling of the Gateway Effect. Health Economics: Vol.6: pp. 521-524.