Essays in Applied Economics: Applications of Transformed Ordinal Quantile Regression

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Okila R. Elboeva

August 2016

Dedication

This dissertation is dedicated to my loving father, Rashid Elboev, who taught me the importance of hard work, diligence, and perseverance.

Acknowledgements

It is my pleasure to express my appreciation to all those who helped and supported me throughout my doctoral studies. First and foremost, I would like to express deepest appreciation to my advisor, Dr. Luiz Lima, for his encouragement, support, and guidance during the research process. His immense knowledge has been an inspiration for me. I would also like to thank Dr. Celeste Carruthers for her continued personal and professional support. Her kindness and diligence defines an exemplary advisor and I am honored to work under her tutelage. I am also greatly appreciative to Dr. William Neilson for lending his expertise, providing guidance, and challenging to do my best. I would like to extend my gratitude to Dr. Danielle Atkins for her willingness to serve and her feedback. Sincerest of thanks to Dr. Don Clark and Dr. Georg Schaur for giving me the opportunity to learn from them and aspire to their work ethics. Lastly, I would like to thank my husband, Shukhrat Musinov, and my children Layla and Elnur for their love and support.

It is because each of you that this accomplishment is possible, thank you.

Abstract

This dissertation consists of three essays on the application of Transformed Ordinal Quantile Regression (TORQUE) developed by Hong and He (2010). TORQUE is based on jittered response, a nonparametric link function, a semiparametric quantile estimation. When the response variable is categorical an application of the standard quantile regression is not optimal. TORQUE technique generalizes ordinary quantile regression, and as a semiparametric method it is more robust than Maximum Likelihood Estimators.

In the first essay I estimate conditional quantiles of happiness using the data from British Household Panel Survey (BHPS) for 2006. I find the continuity assumption of happiness ranking does not hold in this framework, implying the direct application of standard quantile regression could produce biased estimators. Results indicate that income, health, and social factors are very important across all quantiles but decreasing in their magnitude. Education has a significant negative association with happiness at upper quantiles and that females are generally more happier than their male counterparts.

The second essay tests an augmented quantity-quality model of fertility. I focus on the effect of Rosenwald schools on conditional quantiles of fertility for rural black women. Results roughly confirm the model. At the extensive margin, a better access to education increased the probability of having a child from 3.3 to 4.2 percent. I do not find significant effect along the extensive margin. However, OLS estimates infer large and significant negative effects. I also test the same theoretical model for a

sample of women who could have attended Rosenwald schools themselves. I find that school exposure decreased the probability of having a child. Results confirm model predictions.

The third essay examines the relationship between state medical marijuana laws and marijuana consumption among high school students. Unlike other papers we focus on the frequency of marijuana use rather than only on participation. Using frequency data allows to understand the relationship between MMLs and marijuana use for different demographics, such as light smokers vs. heavy smokers. Results imply that MMLs reduce the probability of smoking. This finding is consistent across different groups and estimators.

Table of Contents

List of Tables

List of Figures

Chapter 1

An Analysis of the Distribution of Happiness

1.1 Introduction

Research on human happiness[1] has become a popular topic in recent years. Newer and more comprehensive data made it possible to answer more questions. To conduct welfare evaluations it is important to know the efficiency of a policy tool across the distribution of happiness because the effects of the determinants of happiness are not homogeneous across the conditional distribution. But the common issue with happiness research lies in the characteristics of the data. Most widely available data comes from surveys, where happiness is classified into ordered values with no cardinal interpretation. Although ordered probits and ordered logits can handle this type of data fairly well, they are not robust against distributional misspecification. In addition, ordered probits and ordered logits rely on parallel the regression assumption, in other words, they assume the relationship between each pair of outcomes is the same and therefore may not be very valuable for welfare analysis.[2]

[1] Happiness, subjective well being, and a life satisfaction are used interchangeably throughout this paper.

[2] Also note, that ordered probit produces estimates in terms of predicted probabilities and marginal effects depend on the value of the predictor.

In this paper I estimate conditional quantiles of happiness based on Hong and He's (2010) transformed ordinal quantile regression (TORQUE) method using data from the British Household Panel Survey (BHPS) for 2006. TORQUE is based on jittered (i.e random noise added) response, a nonparametric link function, and a semiparametric quantile estimation. An application of the standard quantile regression is not optimal when the response variable is not continuous and does not have cardinal interpretation. By assuming continuity we are saying that the distance between "very happy" and "happy" is the same as the distance between "very miserable" and "miserable". The transformation of the response variable in TORQUE allows the distance between categories of happiness to vary. The transformed link function is estimated with nonparametric rank based estimator. TORQUE technique generalizes ordinary quantile regression, and as a semiparametric method it is more robust than MLEs because does not rely on any parametric assumption of the conditional distribution. As far as I am are aware, this work is the first to introduce TORQUE methodology in the field of economics.

Research on human happiness can be divided into two main streams. The first stream assumes the happiness survey data has an ordinal interpretation.[3] The most commonly applied estimation technique in this case is ordered logits or probits. As noted earlier, consistency of these estimators hinges on rather strong parametric assumptions of the conditional distribution. The other line of research assumes cardinality of happiness surveys.[4] When cardinality is assumed, the analysis is done by classical ordinary least squares or similar methodology. OLS is very attractive from a practitioner's viewpoint because of its simplicity and relative robustness properties. Using OLS and fixed effect ordered logits, A.Ferrer-i-Carbonell et.al (2004) find that the assumption of ordinality or cardinality of happiness scores makes little difference

[3]Blanchflower et.al (2004) study well being over time in Great Britain and USA using logits, Clark et.al (1996) use ordered probits to test the relationship between relative income and happiness, L.Winkelmann et.al (1998) test non-pecuniary costs of unemployment on life satisfaction.

[4]L.Bruni et.al (2008) investigates the role of relational goods on happiness using World Values Survey, Oswald et al. (2008) use longitudinal data of BHPS to study partial hedonic adaptation,

to the results. This conclusion is based on the comparison of the direction, statistical significance and relative contribution of the estimators. However, the consistency of ordered logit estimators relies on the normality of the error distribution. Further, there is an empirical evidence that the distribution of happiness is skewed. Given the fact that OLS estimators can be inefficient with non-normally distributed errors, these estimators may not be optimal for happiness analysis. In addition, OLS does not provide a comprehensive picture of the underlying relationship between happiness and the variables of interest: OLS just makes predictions of well being of the average person. On the other hand, quantile regression estimation provides richer description of the data, allowing researchers to analyze the effect of the covariate along the entire distribution of the response, rather than just the conditional mean.

This paper is closely related to the work of M.Binder et.al (2011) in using quantile regression to determine the relationship between happiness and explanatory variables for the same dataset. By directly applying Koenker and Basset's (1978) quantile regression M.Binder et.al (2011) implicitly assume that happiness is a continuous variable. My results reveal that the distance between happiness levels is not equal which means that the continuity assumption of the dependent variable does not hold in this application. The largest disparities occur at the tails of the conditional distribution. The comparison between TORQUE and ordinary quantile regression estimators shows that quantile regression can lead to overestimation at lower quantiles and underestimation at upper quantiles.

Results of this paper indicate that health, social life, and marital status are very important factors for happiness. Further, the amount of happiness bought by extra income diminishes with happiness levels. This finding partly confirms the Easterlin Paradox: money doesn't buy happiness. This effect holds for other explanatory variables as well, possibly implying that external factors are not very important for the happiest people and that they are happy just by their nature.

The rest of the paper is organized as follows. Section 1.2 develops the TORQUE methodology. Section 1.3 describes the data and explains variables chosen for the

analysis. Section 1.4 presents results and comparison of TORQUE with OLS, ordered probit and ordinary quantile regression estimators. Section 1.5 concludes and discusses implications of the results and the potential for future research.

1.2 Model

To estimate conditional quantiles of well being, I use Hong and He's (2010) TORQUE methodology. The method is based on the transformed quantile regression model with jittered response and works as follows:

1. Create a jittered response $\widetilde{Y}$ by adding a random noise U$\sim U[0,1)$ to the ordinal random variable Y, so that $\widetilde{Y}_i = Y_i + U_i$. Jittering allows to specify the distributional relationship of the model, which is necessary to identify conditional quantiles of Y.

2. Consider the estimation of the following transformed model:

$$\Lambda(\widetilde{Y}) = X^T\beta + \epsilon \tag{1.1}$$

   where Λ is a monotone function, X is a vector of explanatory variables, β is the vector of linear coefficients, and ϵ is unobserved error with distribution F. This model transformation not only reduces biases due to heteroskedasticity, non-additivity, and non-normality of errors, but it also allows the distance between categories of the ordered response to vary. Different forms of Λ are associated with different econometric and statistical models. Common versions of transformation models include log-linear models, Box-Cox regression models, and proportional hazards models. Most estimation procedures require a parametric assumption for Λ and F and any misspecification can lead to invalid statistical inference.

Recent contributions make it possible to estimate (1.1) when neither Λ nor F is assumed to belong to a known parametric family of functions. For example, Horowitz (1996) proposes a two-step semiparametric method for estimation of transformation models. In addition to being computationally intense, Horowitz's method becomes unreliable when Y moves away from the center of the distribution. This poses a problem when inference across all of the conditional distribution is of particular interest. Chen (2002), on the other hand, develops a rank-based estimator for Λ given a consistent estimator for β. The general idea of the rank estimator is as follows: if Λ is nondecreasing function, then $\Lambda_i > \Lambda_j$ implies that

$$P(X_i^T\beta > X_j^T\beta) > P(X_j^T\beta > X_i^T\beta).$$

This suggests estimating Λ so as to make the ordering of $\hat{\Lambda}$ as close as possible to the ordering of $X^T\beta$. TORQUE methodology uses Chen's estimator to estimate the transformation function Λ. Like any rank-based estimator, Chen's $\hat{\Lambda}$ is highly efficient, robust to outliers and easy to implement.

Note that (1.1) continues to hold if Λ, β, and ϵ are replaced with $a\Lambda$, $a\beta$, and $a\epsilon$ for any positive constant a, or with $\Lambda + c$, $\beta + c$, and $\epsilon + c$ for any constant c. Location and scale normalization is needed for identification. Location normalization is achieved by setting $\Lambda(\tilde{y}_0) = 0$ for some $\tilde{y}_0$, where $\tilde{y}_0$ is the value of $\tilde{y}$ when $\Lambda = 0$. For scale normalization the first coefficient of β is set to one. The choice of uniform distribution for U is for convenience only, other types of distributions lead to a different $\Lambda(\widetilde{Y})$ function without changing its distribution.

The consistency of $\hat{\Lambda}$ in (1.1) requires an initial consistent estimator b of β. Depending on the model assumptions, one can use least squares estimator from regressing $\widetilde{Y}_i$ on X_i.

3. Using the same estimation method as in Chen (2002), obtain an estimate of Λ for each $\widetilde{y}$ as:

$$\hat{\Lambda}(\tilde{y}) = \underset{\Lambda \in M_\Gamma}{\arg\max} \left\{ \Gamma(\tilde{y}, \Lambda, b) = \sum_{i \neq j} (d_{i\tilde{y}} - d_{j\tilde{y_0}}) * 1\{X_i^T b - X_j^T b \geq \Lambda\} \right\} \quad (1.2)$$

where $d_{i\tilde{y}} = 1\{\widetilde{Y}_i \geq \tilde{y}\}$, and $d_{j\tilde{y_0}} = 1\{\widetilde{Y}_j \geq \tilde{y_0}\}$ for some $\tilde{y_0}$ chosen under location normalization assumption of $\Lambda(\tilde{y_0}) = 0$, and M_Γ is a prechosen compact set in R.

4. Consider to estimate τth conditional quantile of Λ given $X = x$:

$$Q_\tau(\Lambda(\widetilde{Y})|X) = \alpha_\tau + X^T \beta_\tau$$

where $\tau \in (0, 1)$, $\alpha_\tau \in R$, $\beta \in R^p$, and $Q_\tau(\Lambda)$ denotes τth quantile of Λ. Quantile estimates of $(\alpha_\tau, \beta_\tau)$ can be calculated by solving:

$$(\hat{\alpha}_\tau, \hat{\beta}_\tau) = \underset{\alpha \in R \ \beta \in R^p}{\arg\min} \sum_{i=1}^{n} \rho_\tau(\hat{\Lambda}(\widetilde{Y}_i) - \alpha - X_i^T \beta) \quad (1.3)$$

where $\rho_\tau(r) = (\tau I(r > 0) + (1 - \tau) I(r < 0))|r|$ is the Koenker and Basset's quantile loss function.

If the predicted quantiles of the original ordinal response, denoted by $Q_\tau(Y|X)$, is of particular interest, it is easy to show that quantile regression's invariance to monotone transformation property allows for $Q_\tau(Y|X)$ to be recovered:

$$Q_\tau(Y|X) = \lfloor \Lambda^{-1}(Q_\tau(\Lambda(\widetilde{Y}_i)|X) \rfloor \quad (1.4)$$

with $\lfloor \cdot \rfloor$ being the integer part of nonnegative number.

1.3 Data

To predict happiness I use the data from British Household Panel Survey (BHPS) for 2006. BHPS is a household-based survey dataset following every adult member of sampled households from the United Kingdom. I proxy individual's subjective well being for happiness. The way in which subjective well being is measured is similar to that used in other happiness surveys. It reflects answers to the question "How dissatisfied or satisfied are you with....your life overall?" on a seven point Likert scale ranging from "Not satisfied at all" (1) to "Completely satisfied" (7). The distribution of happiness (Figure 1.1) is quite skewed and very similar to that of Binder's (2011), which justifies the use of quantile regression.

There are two potential issues in using self-reported subjective well being as a measure for happiness: a measurement problem and validity concerns. A measurement problem arises because survey answers can be highly variable and individuals' judgments can be affected by transient factors such as mood at the time of

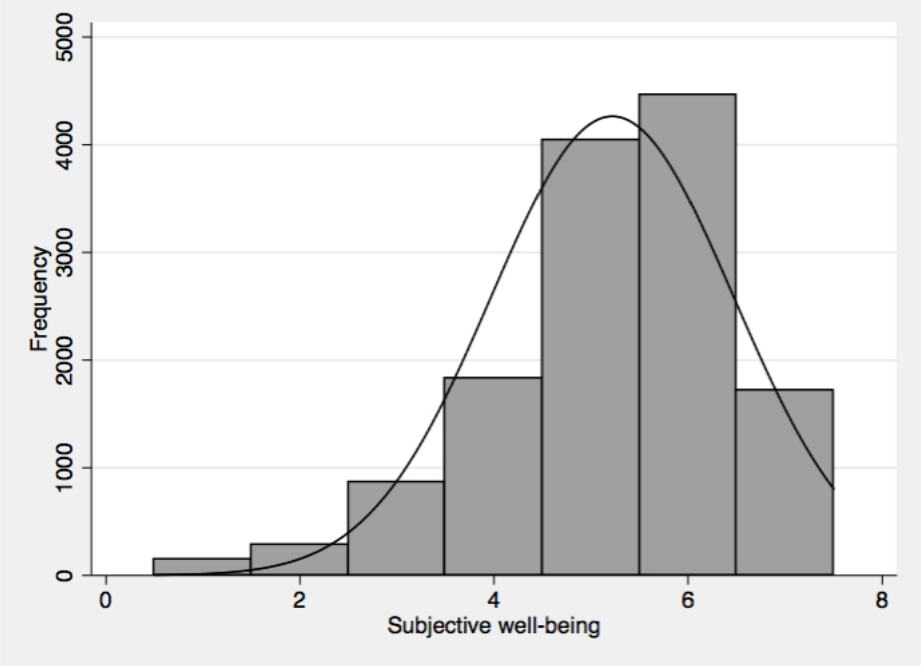

Figure 1.1: The distribution of subjective well being overlaid with normal-density plot

the survey. However, Krueger et.al (2008) find that the estimated degree of reliability of the subjective well being data is high enough to support research on happiness. The validity hinges on the assumption that individuals are willing and able to order their well being in a way that is invariant to a positive monotone transformation. The general view is that although subjective measurements are not perfect, they do reflect substantive feelings and participants have little trouble ordering and assessing them (Easterlin, 1974,2001). For the BHPS 2006 only 4.27 percent of the information on life satisfaction was unavailable due to missing data or individuals not answering the question.

1.3.1 Predictors

In psychology, a notion exists that utility recovers after an adverse event and that happiness is determined by quite stable personality traits. Psychologists call this adjustment process "hedonic adaptation". Complete hedonic adaptation implies that life circumstances play negligible role on well being and there is little that public policies can do to improve well being. In economics, on the other hand, it is generally assumed that changes in well being are permanent with no adaptation. However, a wide literature provides theoretical and empirical evidence that events such as marriage, birth of children, retirement, income, loss of a job, etc. permanently affect subjective well being with some degree of adjustment. For example, Oswald and Powdthawee (2008) present evidence of hedonic adaptation to disability. They estimate that the extent of adaptation for moderate and severe disability is within 30-50 percent. Clark and Oswald (1994) show that there is only a partial adaptation to long-term unemployment. Easterlin (2001, 2003, 2005) argues that people do not return to their baseline happiness level after major events and that happiness is a variable determined by predictors.

In choosing explanatory variables I closely follow Binder (2011). Specifically, I analyze how happiness is correlated with income, health and social factors, combined

Table 1.1: Summary statistics from BHPS

			Quantiles				
Variable name	Mean	St. Dev	Min	.25	Mdn	.75	Max
Subjective well being	5.22	1.25	1	5	5	6	7
Logarithm of income	10.04	0.68	−0.51	9.68	10.08	10.46	12.50
Health factor Index	0.00	1.34	−6.89	−0.66	0.48	0.99	1.63
Social factor Index	0.00	1.12	−5.82	−0.70	0.12	0.79	1.86
Age	45.88	18.52	15.00	31.00	44.00	60.00	99.00
Age_Sq	342.79	373.34	0.01	47.38	199.27	534.37	2821.35
Education	3.23	1.76	1.00	1.00	3.00	4.00	7.00
Number of children	0.49	0.91	0.00	0.00	0.00	1.00	7.00

Number of observations = 13348. (Dummies for gender, marital and job statuses are omitted)

with controls for employment, marital status, age, gender, number of children, and education. Table 1.1 provides summary statistics.

One of the main variables of interest is income. The relationship between income and happiness is a complex one. Most studies generally suggest positive but diminishing returns to income. Binder results support the idea of positive correlation between income and happiness but only at lower quantiles of happiness. He finds that income does not matter for people on the upper part of the distribution. Easterlin (1995) reports a positive correlation in a cross-section of people, but the effect is only transitory. Using panel data for Russia, Graham et.al (2004) finds a reverse causation, that is happier people are more likely to report higher future incomes.

To measure income, I use the logarithm of net equalized annual household income in British pounds before housing costs using McClements equivalence scale.

Table 1.2: Average incomes at happiness categories

	Happiness categories							
	1	2	3	4	5	6	7	All categories
Ave Income in British pounds	19109.75	20901.35	24042.49	25038.39	28787.58	30539.74	24922.44	27786.06
Ave Income in logs	9.69	9.77	9.91	9.95	10.08	10.14	9.93	10.04

Equivalence scales take into account not only higher spending needs for larger households but also economies of scale in number of family members that live together. Data shows (Table 1.2) that the average income for all categories is increasing except for the happiest people. On average, the income level of the happiest person is 10.3% lower than the whole sample average. Generally, however, there is a positive correlation between income and happiness (Table 1.3).

The next covariate is the Health Index. The Index includes individual's subjective and objective measures of health. Subjective measure includes a self assessment of health for the last 12 months, which is scaled on a five point Likert scale, ranging from "poor" (1) to "excellent"(5). While the causality between subjective

Table 1.3: Correlation between happiness and income, health, and social factors

	Happiness	Income	Health Index	Social Index	Education	Age
Happiness	1.0000					
Income	0.0786	1.0000				
Health Index	0.2743	0.1312	1.0000			
Social Index	0.3861	−0.0595	0.0723	1.0000		
Education	0.0121	0.2906	0.1919	−0.1268	1.0000	
Age	0.0623	−0.0241	−0.1910	0.0438	−0.3058	1.0000

Measurements are obtained after cleaning the BHPS dataset. Income is measured as a logarithm of net household income before housing costs and adjusted using McClements equivalence scale.

assessment of health and happiness can run both ways, objective health most likely determines happiness. For an objective measure of health I include number of days spent in hospital, number of visits to a general practitioner, and number of serious accidents in the previous year. I aggregate subjective and objective measures of health using Principal Component Analysis (PCA). PCA allows me to represent information contained in health variables with one continuous variable while removing idiosyncratic noise. The goodness of fit of the Health Index (Kaiser-Meyer-Olkin measure is 0.63) is acceptable and accounts for ρ=45 percent of the indicators' variance.

The next variable of interest is a Social Factor Index which represents the quality of one's social life. Similar to the Health Index, it is summarized by subjective and objective measures. Individual's own assessment is measured on seven-point Likert scale from "not satisfied at all"(1) to "completely satisfied"(7) and reflects responses to a question "How dissatisfied or satisfied are you with your social life?". The objective component of the Index is measured by the "frequency of talking to neighbors," and "frequency of meeting people," ranging from 1-5. This Index is also computed by PCA, and the goodness of fit is acceptable (Kaiser-Meyer-Olkin measure is 0.5464) and accounts for ρ=41.73 percent of the indicators' variance.

I also include controls consisting of employment dummies, marital status dummies, education dummies, number of children, gender, age, and age^2 (the squared difference between age and mean age instead of using age squared). Reverse causation between marital status and happiness is possible because happier people are more likely to get married and stay married. However, Winkelmann and Winkelmann (1998) find that marriage effects double when differencing out individual fixed effects using longitudinal data, so marriage seems to protect mental well being. Findings about the relationship between age and happiness are not unanimous: while Binder (2011) finds positive to non-significant association between age and happiness, other studies (Blanchflower & Oswald, 2004) find a U-shaped relationship.

1.4 Results

1.4.1 The Analysis of the Link Function Λ

The consistency of $\hat{\Lambda}(\tilde{y})$ depends on the consistency of the initial estimate of β. I use the median regression estimator b because it is more robust to outliers.

Figure 1.2(a) shows the plot of the estimated function of $\Lambda(\tilde{y})$. At the mean values the relationship between $\hat{\Lambda}$ and $\tilde{Y}$ is very close to linear. In contrast, at extreme values the slopes deviate from the fitted line. A flatter(steeper) slope of $\hat{\Lambda}(\tilde{y})$ indicates that the estimated distance between levels of well being is shorter(longer) than the ones obtained by assuming the continuity of the well being. Based on the above result, I conclude the continuity assumption of the subjective well being does not hold. This conclusion means that the direct application of the quantile regression most likely

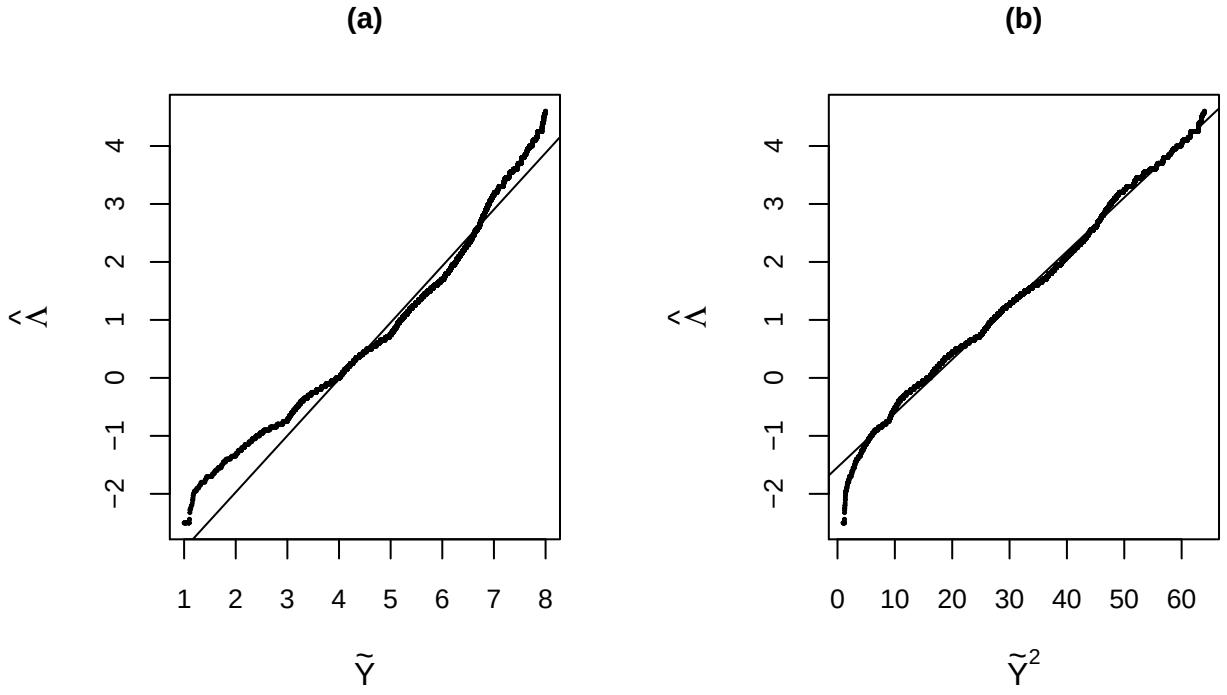

Figure 1.2: Estimation results for Lambda-Happiness

leads to an attenuation bias away from zero at lower quantiles and an attenuation bias toward zero at upper quantiles of the conditional distribution.

TORQUE estimators from step four explain the change in Λ given a marginal change in predictors for a given quantile: $\frac{\delta\Lambda_\tau}{\delta x}$. However, we mainly care about marginal effects on Y. Quantile regression's invariance to monotone transformation property allows to recover $\frac{\delta Y_\tau}{\delta x}$. From Figure 1.2(a) it is obvious that the association between $\hat{\Lambda}$ and $\tilde{Y}$ resembles a quadratic relationship. The regression of $\hat{\Lambda}$ on $\tilde{Y}^2$ shows a very good fit except at the extreme left tail (Figure 1.2(b)). Since the estimated fraction of observations [5] that fall into that tail is about 0.4 percent, using quadratic relationship between $\hat{\Lambda}$ on $\tilde{Y}^2$ to rescale $\hat{\beta}$ should cause very little bias.

1.4.2 Main Results

Main findings are given in Table 1.4. I report β_τ estimates from TORQUE methodology as well as estimates from untransformed ordinary quantile regression (QR) for the 10, 25, 50, 75 and 90 quantiles of the conditional distribution. TORQUE estimators are recovered in the following way. First, I regress $\hat{\Lambda}_i$ on $\tilde{Y}_i^2$: $\hat{\Lambda}_i = a + c\tilde{Y}_i^2$. Next, I rearrange to get: $\tilde{Y} = \left(\frac{\hat{\Lambda}}{\hat{c}} - \frac{\hat{a}}{\hat{c}}\right)^{\frac{1}{2}}$. Finally, $\frac{\delta Y_\tau}{\delta x} = \hat{\beta}_\tau \frac{1}{2\hat{c}} \left(\frac{\hat{\Lambda}_\tau}{\hat{c}} - \frac{\hat{a}}{\hat{c}}\right)^{-\frac{1}{2}}$, where $\hat{\Lambda}_\tau$ is the average of the fitted value at a given quantile. Fitted values and $\hat{\beta}_\tau$ are obtained from step four of the model.

A first observation is that marginal effects are heterogeneous across quantiles and they are mostly decreasing in levels of well being. This result confirms previous findings. Log of equalized income has positive but diminishing association with well being and statistically significant at the five percent level for the 10, 25, 50, 75 quantiles. An interesting finding is that results for the happiest ten percent of the population sample do not support the idea that income buys happiness. For a given sample, decreasing income marginal effects cannot be explained by diminishing marginal utility of income because on average, the happiest people are poorer than

[5] I used $\tilde{Y} < 1.3$, the fraction of observations for $\tilde{Y} < 1.5$ is only 0.54 percent.

Table 1.4: TORQUE model of estimated happiness results

	Dependent variable: Subjective well being									
	β^{TORQUE} (q10)	β^{QR} (q10)	β^{TORQUE} (q25)	β^{QR} (q25)	β^{TORQUE} (q50)	β^{QR} (q50)	β^{TORQUE} (q75)	β^{QR} (q75)	β^{TORQUE} (q90)	β^{QR} (q90)
logincome	0.1499***	0.1622***	0.1093***	0.1283***	0.0892***	0.1139***	0.0453***	0.0367*	0.0253	0.0223
	(0.0339)	(0.0385)	(0.0222)	(0.0233)	(0.0172)	(0.0209)	(0.0174)	(0.0192)	(0.0225)	(0.0190)
Social Index	0.4636***	0.5053***	0.3956***	0.4575***	0.3530***	0.3861***	0.3311***	0.2860***	0.2539***	0.2606***
	(0.0201)	(0.0176)	(0.0111)	(0.0153)	(0.0101)	(0.0089)	(0.0099)	(0.0157)	(0.0129)	(0.0111)
Health Index	0.3012***	0.3340***	0.2282***	0.2655***	0.189***	0.2156***	0.1601***	0.1521***	0.1205***	0.1057***
	(0.0212)	(0.0198)	(0.0147)	(0.0137)	(0.0103)	(0.0117)	(0.008)	(0.0111)	(0.0122)	(0.0113)
education	0.0248*	0.0427***	0.0334***	0.0439***	0.0139**	0.0164***	−0.0259***	−0.0121**	−0.0389***	−0.0525***
	(0.0132)	(0.0132)	(0.0078)	(0.0083)	(0.0071)	(0.0062)	(0.0067)	(0.0061)	(0.0085)	(0.0078)
female	−0.0305	−0.0532	−0.0173	−0.0142	0.0008	0.0104	0.0620***	0.0442**	0.0551*	0.0521*
	(0.0345)	(0.0419)	(0.0242)	(0.0282)	(0.0251)	(0.0214)	(0.0271)	(0.0192)	(0.0309)	(0.0270)
age	0.0017	0.0016	0.0015	0.0008	0.002*	0.0019	0.0023*	0.0025**	0.0008	0.0014
	(0.0022)	(0.0020)	(0.0012)	(0.0015)	(0.0012)	(0.0012)	(0.0016)	(0.0012)	(0.0014)	(0.0015)
age_sq	0.0002***	0.0002***	0.0003***	0.0003***	0.0003***	0.0003***	0.0003***	0.0003***	0.0002***	0.0003***
	(0.0001)	(0.0001)	(0.00004)	(0.0001)	(0.00005)	(0.0001)	(0.0001)	(0.00004)	(0.0001)	(0.00005)
number of children	−0.0282	−0.0450*	−0.0625***	−0.0619***	−0.0522***	−0.0381**	−0.0458***	−0.0315**	−0.0515***	−0.0650***
	(0.0272)	(0.0249)	(0.0171)	(0.0163)	(0.0103)	(0.0144)	(0.0204)	(0.0124)	(0.0216)	(0.0160)
Marital status dummies (reference group: married)										
separated	−0.4395***	−0.4217**	−0.4131***	−0.4059***	−0.3036***	−0.3484***	−0.3405***	−0.3026***	−0.2109	−0.2747*
	(0.1657)	(0.1659)	(0.1138)	(0.1527)	(0.0752)	(0.0967)	(0.0959)	(0.0739)	(0.1668)	(0.1496)
divorced	−0.3040***	−0.3530***	−0.3726***	−0.4633***	−0.3308***	−0.3644***	−0.2796***	−0.2314***	−0.1364**	−0.1395**
	(0.061)	(0.0722)	(0.0629)	(0.0609)	(0.0436)	(0.0461)	(0.0508)	(0.0396)	(0.0614)	(0.0581)
widowed	−0.3758***	−0.3178***	−0.3496***	−0.3637***	−0.3269***	−0.3248***	−0.2369***	−0.1511***	−0.1405*	−0.1753***
	(0.0875)	(0.0937)	(0.0639)	(0.0672)	(0.0475)	(0.0568)	(0.0548)	(0.0581)	(0.0749)	(0.0446)
never married	−0.2495***	−0.2272***	−0.2514***	−0.2973***	−0.2376***	−0.2689***	−0.1589***	−0.1221***	−0.1368***	−0.1283***
	(0.0687)	(0.0606)	(0.0391)	(0.0386)	(0.0305)	(0.0399)	(0.0384)	(0.0304)	(0.0369)	(0.0454)
in a civil partnership	−0.0991	−0.0363	−0.1676	−0.2339	0.0902	0.1421	0.1501	−0.0566	−0.0798	−0.1624
	(0.1865)	(0.2334)	(0.1632)	(0.1833)	(0.0773)	(0.1219)	(0.1503)	(0.0555)	(0.0795)	(0.1614)
Employment status dummies (reference group: self-employed)										
employed	−0.1033	−0.0929	−0.0262	−0.0624	−0.0724*	−0.0641*	0.0013	−0.0493	0.0306	0.0349
	(0.0839)	(0.0708)	(0.0429)	(0.0491)	(0.0398)	(0.0371)	(0.0503)	(0.0344)	(0.0417)	(0.0506)
unemployed	−0.8203***	−0.9963***	−0.6181***	−0.7909***	−0.4671***	−0.5606***	−0.2975***	−0.3363***	0.0286	0.0886
	(0.1541)	(0.1626)	(0.0822)	(0.0952)	(0.0935)	(0.1006)	(0.1293)	(0.0783)	(0.1247)	(0.1204)
retired	−0.0669	−0.0363	0.0497	0.0804	0.1*	0.0786	0.1674***	0.1729***	0.1957***	0.1432**
	(0.1311)	(0.1002)	(0.0668)	(0.0729)	(0.0605)	(0.0557)	(0.0547)	(0.0547)	(0.0493)	(0.0630)
maternity leave	1.0875***	0.8856***	0.8664***	0.8781***	0.7069***	0.6320***	0.6609***	0.5176***	0.7258***	0.6282***
	(0.1735)	(0.2593)	(0.2001)	(0.1765)	(0.1454)	(0.1453)	(0.1363)	(0.1248)	(0.2203)	(0.2231)
family care	−0.2184**	−0.1051	−0.1258	−0.1770**	−0.0797	−0.1370**	0.0014	−0.0339	0.0901	0.0979
	(0.104)	(0.1106)	(0.0948)	(0.0679)	(0.0706)	(0.0664)	(0.0541)	(0.0596)	(0.0824)	(0.0715)
full time student, school	−0.1137	−0.1660	0.0014	0.0331	−0.0151	0.0317	0.0058	−0.0159	0.1012	0.0379
	(0.1196)	(0.1089)	(0.0514)	(0.0716)	(0.09)	(0.0614)	(0.0761)	(0.0660)	(0.0763)	(0.0716)
sick, disabld	−0.777***	−0.7913***	−0.4973***	−0.7110***	−0.4572***	−0.7081***	−0.4149***	−0.5665***	−0.2287**	−0.1143
	(0.1593)	(0.1647)	(0.0864)	(0.1007)	(0.0827)	(0.0846)	(0.0665)	(0.1003)	(0.1061)	(0.1279)
Pseudo-R squared	0.1605	0.2057	0.1459	0.1735	0.1225	0.1355	0.1253	0.0426	0.1075	0.1014
Observations	13,348	13,348	13,348	13,348	13,348	13,348	13,348	13,348	13,348	13,348

Note: Standard errors for TORQUE estimators are obtained with Delta method, quantile regression standard errors are bootstrapped. *p<0.1; **p<0.05; ***p<0.01

the average person. This finding confirms Easterlin Paradox that more money does not make people happier. But this is true for only very happy people.

Not surprisingly, health and social factors' coefficients are positive and statistically very significant. While the effect is decreasing with quantiles for both, the social factor is twice as important for happiness as health. There is a possibility that the Health Index is endogenous because happier people are less likely to suffer from stress related diseases. As described earlier, the Health Index is constructed using subjective and objective assessments of health. The subjective assessment reflects an individual's report on mental well being and is more likely to be the source of endogeneity. I also repeat the analysis with the Health Index excluding subjective well being component.

Marginal effects for education are highly significant across the distribution. The magnitude of the effect decreases in quantiles of happiness and for the top twenty-fifth quantile it becomes negative. Contrary to what we observe with incomes, average education levels are decreasing with happiness until the top 10 percent of the happiest people. The top ten percent of happiest people's average education level is higher than sample's average education level but less than that of most miserable people. Further, the relative contribution of education to happiness is at least ten times less in magnitudes than health and social factors. This could be associated with the fact that education is highly correlated with income, health, etc., and if the causality runs from education to these factors, then controlling for these factors underestimates education effects.

Females are generally happier than their male counterparts; however, the difference in not very large. Age has very little and mostly statistically insignificant effect. I do not find the popular U-shaped relationship between age and happiness. Additional children decrease happiness, but this relative contribution is much smaller compared to the most other controls' effects.

Being married seems to be the best option for marital status. Divorce, separation or death of the spouse significantly reduces well being. In addition, results indicate

that a married life is superior to a bachelor lifestyle. The positive relationship between marriage and happiness can be attributed to the protection and selection effects of a marriage. Individuals in a civil partnership are less happy than married people, although the effect is not significant. The explanation for this happiness gap could be the incomplete institutionalization of cohabitation. Overall, the relative importance of these factors is comparable to the magnitude of quality of social life and the health factor effects.

Having a job is highly associated with better well being, regardless of if one is one's own boss or works for others, except for those who are in the medium of the distribution. For them, being self-employed seems to increase well being. Being dick and disabled is associated with significantly lower happiness levels: its relative contribution is the highest of all predictors. Surprisingly, maternity leave increases the happiness of females.[6]

Now I compare estimates from TORQUE and untransformed QR. The statistical significance of the coefficients is almost identical in both methods. In contrast, their levels show large disparities. From the relationship of $\hat{\Lambda}$ and $\tilde{Y}$ (Figure 1.2) it is expected that, on average, QR coefficients should overestimate at lower quantiles and underestimate at higher quantiles of happiness. Estimation results confirm this expectation. For example, tenth and twenty-fifth quantiles of $\hat{\beta}^{TORQUE}_{income}$ are equal to 0.1499 and 0.10933, whereas $\hat{\beta}^{QR}_{income}$ are equal to 0.1622 and 0.1283, which equals to overestimation of 8 and 17.4 percent. Marginal effects for health, social indexes, education and gender at the tenth and twenty-fifth quantile are overestimated by 9-74 percent and 16-31 percent, respectively. The overestimation of marginal effects for marital status dummies is more consistent at the twenty-fifth quantile, ranging between 4-24 percent. Further analysis reveals that upper quantile marginal effects obtained by ordinary quantile regression are typically lower than TORQUE estimates. For example, seventy fifth quantile $\hat{\beta}^{QR}$ for health and social indexes, education,

[6]There is only one observation where a male is on maternity leave.

Table 1.5: Happiness prediction performance measures for TORQUE, QR and OLS

	q10		q25		q50		q75		q90		
	TORQUE	*QR*	*TORQUE*	*QR*	*TORQUE*	*QR*	*TORQUE*	*QR*	*TORQUE*	*QR*	*OLS*
Mean Absolute Errors	0.744	1.518	0.615	1.028	0.495	0.834	0.392	1.004	0.371	1.346	0.839
Difference in:											
levels		−0.774		−0.413		−0.339		−0.612		−0.975	−0.344
percentages		−51		−40.2		−40.6		−61		−62.5	−41

MAE from Quantile Regression is taken as the base for calculation of differences. MAE from OLS is compared with MAE TORQUE at fiftieth quantile.

gender are lower by 5-53 percent than $\hat{\beta}^{TORQUE}$. However, results do not show this consistency at the top tenth quantile.

Table 1.5 reports mean absolute errors between predicted responses and the observed errors for TORQUE, QR, and OLS methods. Since TORQUE cannot identify location, estimates of constants for each quantile from QR are used instead. TORQUE outperforms in all cases with its MEA being from forty to sixty percent less than MAE from QR. As expected, the biggest difference is at the tails of the distribution. OLS performance is similar to QR at the fiftieth quantiles.

It is very common to use ordinary least squares and ordered probits for this type of analysis. Table 1.6 provides estimates from TORQUE and ordered probits. Ordered probit estimates are obtained at the means of the independent variables. For comparison, in all cases the coefficient of the first predictor is set to one. The direction of coefficients, their statistical significance and relative contributions are very close for $\hat{\beta}_{0.5}^{TORQUE}$ and $\hat{\beta}^{OP}$, except for variables that explain the most variation of the response. However, results suggest that TORQUE estimators and ordered probit estimators may tell different stories at the tails of the distribution. Note that ordered probits can be sensitive to distributional misspecification, although in this case the bias does not seem to be large.

For further analysis I also compare $\hat{\beta}^{OLS}$ with $\hat{\beta}_{0.5}^{TORQUE}$. Results (Table 1.7) show that OLS performs very well: statistical significance and the direction of the effects is almost identical to that of TORQUE's . Although $\hat{\beta}^{OLS}$ are larger than

Table 1.6: ORDERED PROBITS model of estimated happiness results

	Dependent variable: subjective well being			
	OP	*TORQUE*		
		(q25)	(q50)	(q75)
logincome	1***	1***	1***	1***
	(0.0017)	(0.0222)	(0.0172)	(0.0174)
Social Index	4.4302***	3.6194***	3.9574***	7.3091***
	(0.0017)	(0.0111)	(0.0101)	(0.001)
Health Index	2.3545***	2.0878***	2.12***	3.5342***
	(0.0011)	(0.0147)	(0.0103)	(0.008)
education	0.0428	0.3056***	0.1558**	−0.5717***
	(0.0007)	(0.0078)	(0.0071)	(0.0067)
female	0.162	−0.1583	0.009	−0.5717***
	(0.0022)	(0.0242)	(0.0251)	(0.0271)
age	0.0277	0.0146	0.0224*	0.0508*
	(0.0001)	(0.0012)	(0.0012)	(0.0016)
age_sq	0.0346***	0.003***	0.003***	0.007***
	(4.33e-06)	(0.0001)	(0.0001)	(0.0001)
number of children	−0.6267***	−0.5718***	−0.5852***	−1.011***
	(0.0014)	(0.0171)	(0.0103)	(0.0204)
Marital status dummies (reference group: married)				
separated	−3.3077***	−3.7795***	−3.4047***	−7.5166***
	(0.0038)	(0.1138)	(0.0752)	(0.0959)
divorced	−3.1734***	−3.4099***	−3.7081***	−6.1722***
	(0.0028)	(0.0629)	(0.0436)	(0.0508)
widowed	−3.1605***	−3.1985***	−3.6648***	−5.2295***
	(0.0031)	(0.0639)	(0.0475)	(0.0548)
never married	−2.5050***	−2.3***	−2.6648***	−3.5077***
	(0.0033)	(0.0391)	(0.0305)	(0.0384)
in a civil partnership	−0.8370	−1.5333	1.0112	3.3113
	(0.0144)	(0.1632)	(0.07731)	(0.1503)
Employment status dummies (reference group: self-employed)				
employed	−0.5842	−0.2406	−0.8117*	0.0287
	(0.0046)	(0.0429)	(0.0398)	(0.0503)
unemployed	−26.4982***	−5.6556***	−5.2365***	−6.5673***
	(0.0046)	(0.0822)	(0.0935)	(0.1293)
retired	1.0811	0.4538	1.12	3.6953***
	(0.0068)	(0.0668)	(0.0605)	(0.0547)
maternity leave	14.2805***	7.7026***	7.8913***	16.4989***
	(0.03)	(0.2001)	(0.1454)	(0.1363)
family care	−1.0681	−1.151	−0.8935	0.0309
	(0.006)	(0.0948)	(0.0706)	(0.0541)
full time student, school	−0.009	0.0137	−0.15	0.1302
	(0.007)	(0.0514)	(0.09)	(0.0761)
sick, disabled	−2.6442***	−4.5498***	−5.1267***	−9.1589***
	(0.005)	(0.0864)	(0.0827)	(0.0665)
Observations	13,348	13,348	13,348	13,348
R^2	0.0896	0.1459	0.1225	0.1253

Note: Standard errors are obtained prior to scaling *p<0.1; **p<0.05; ***p<0.01

$\hat{\beta}_{0.5}^{TORQUE}$, the relative contribution of the controls is very similar. OLS method is very attractive because of its simplicity and relative robustness properties, but unlike quantile regression analysis it cannot give a complete picture about the effects of variables across the conditional distribution.

As mentioned earlier, it is hard to make a strong causal claim about happiness and the Health Index used in the main analysis. For a robustness check I reestimate the model with a new index by excluding the subjective assessment of health. Table 1.8 reports results with "endogenous" Health Index for a selected group of controls. New results show that now explanatory variables explain more variation in happiness, except for marital status dummies and a dummy for a maternity leave. It appears to be that a dummy for a maternity leave is picking out some of the effects of individuals' assessments of the subjective well being. Marital status dummies' relative contribution is still very large. Overall, using a more robust measure for the Health Index does not alter my main conclusions by much.

1.5 Conclusions

In this paper I bring into the field of economics a new econometric approach for handling ordinal data. While ordered probits and ordered logits perform pretty well for this type of data, their reliance on parametric assumptions make estimation sensitive to any misspecifications. Hong and He's (2010) TORQUE methodology is applied to survey data to learn the causes of subjective well being. I show that "subjective well being" measurement from BHPS for the year 2006 does not have cardinal meaning: the distance between different levels of well being is not equal.

TORQUE methodology is applied to a common set of explanatory variables. Results indicate that health, good social life, employment and marital status are very important determinants of happiness. Economists emphasize the importance of life circumstances to well being, especially income. The implication is that public policy measures aimed at increasing income will lead to an increase of

Table 1.7: OLS model of estimated happiness results

	Dependent variable: subjective well being	
	TORQUE	*OLS*
	(1)	(2)
logincome	0.0892***	0.0980***
	(0.0172)	(0.0154)
Social Index	0.3530***	0.3931***
	(0.0101)	(0.0086)
Health Index	0.1891***	0.2132***
	(0.0103)	(0.0078)
education	0.0139**	0.0122**
	(0.0071)	(0.0060)
female	0.0008	0.0081
	(0.0251)	(0.0198)
page	0.002*	0.0021**
	(0.0012)	(0.0010)
age_sq	0.0003***	0.0003***
	(0.0001)	(0.00004)
number of children	−0.0522***	−0.0525***
	(0.0103)	(0.0122)
Marital status dummies (reference group: married)		
separated	−0.3037***	−0.3624***
	(0.0752)	(0.0684)
divorced	−0.3308***	−0.3245***
	(0.0436)	(0.0356)
widowed	−0.3269***	−0.3065***
	(0.0475)	(0.0433)
never married	−0.2377***	−0.2146***
	(0.0305)	(0.0311)
in a civil partnership	0.0902	−0.0335
	(0.0773)	(0.1143)
Employment status dummies (reference group: self-employed)		
employed	−0.0724*	−0.0572
	(0.0398)	(0.0382)
unemployed	−0.4671***	−0.5559***
	(0.0935)	(0.0658)
retired	0.0999	0.0228
	(0.0605)	(0.0508)
maternity leave	0.7039***	0.7219***
	(0.1454)	(0.1458)
family care	−0.0797	−0.1322**
	(0.0706)	(0.0542)
full time student, school	−0.0134	−0.0230
	(0.09)	(0.0592)
sick, disabled	−0.4573***	−0.6055***
	(0.0827)	(0.0616)
Observations	13,348	13,348
R^2		0.2568
Adjusted R^2	0.1225	0.2555

Note: TORQUE estimates are given for the fiftieth quantile for selected variables. *p<0.1; **p<0.05; ***p<0.01

Table 1.8: TORQUE results with exogeneous Health Index

	Dependent variable: Subjective well being									
	β^{TORQUE}	β^{QR}	β^{TORQUE}	β^{QR}	β^{TORQUE}	β^{QR}	β^{TORQUE}	β^{QR}	β^{TORQUE}	β^{QR}
	(q10)	(q10)	(q25)	(q25)	(q50)	(q50)	(q75)	(q75)	(q90)	(q90)
logincome	0.1769***	0.2084***	0.1298***	0.1436***	0.1011***	0.1284***	0.0545***	0.0367**	0.0398*	0.0393**
	(0.0288)	(0.0416)	(0.0272)	(0.0328)	(0.0164)	(0.0196)	(0.0203)	(0.0167)	(0.0255)	(0.0189)
Social Index	0.5116***	0.5662***	0.4195***	0.4872***	0.3708***	0.3983***	0.3451***	0.2788***	0.2618***	0.2696***
	(0.0177)	(0.0182)	(0.0146)	(0.0152)	(0.0119)	(0.0098)	(0.011)	(0.0174)	(0.0159)	(0.0115)
Health Index	0.1863***	0.2198***	0.14***	0.1657***	0.1201***	0.1381***	0.0982***	0.0882***	0.0668***	0.0617***
	(0.0144)	(0.0248)	(0.0121)	(0.0159)	(0.0085)	(0.0116)	(0.0139)	(0.0112)	(0.012)	(0.0102)
education	0.0392***	0.0479***	0.0433***	0.0584***	0.0258***	0.0268***	−0.0218***	−0.0088	−0.0391***	−0.0470***
	(0.0123)	(0.0123)	(0.0109)	(0.0094)	(0.0079)	(0.0070)	(0.0094)	(0.0059)	(0.0093)	(0.0077)
female	−0.0552	−0.0676	−0.0484**	−0.0208	−0.0184	−0.0111	0.0481**	0.0250	0.0419*	0.0420*
	(0.0451)	(0.0425)	(0.0222)	(0.0272)	(0.022)	(0.0208)	(0.02)	(0.0178)	(0.027)	(0.0255)
age	0.0012	−0.0009	0.0012	0.0019	0.0024*	0.0023**	0.0022	0.0025**	0.0009	0.0018
	(0.0019)	(0.0020)	(0.0013)	(0.0015)	(0.0015)	(0.0011)	(0.0016)	(0.0010)	(0.0017)	(0.0012)
number of children	−0.0317	−0.0514*	−0.0613***	−0.0597***	−0.0359**	−0.0451***	−0.0329**	−0.0249**	−0.0545***	−0.0574***
	(0.0192)	(0.0268)	(0.0189)	(0.0194)	(0.0181)	(0.0154)	(0.0149)	(0.0114)	(0.02)	(0.0156)
Marital status dummies (reference group: married)										
separated	−0.3533***	−0.3797**	−0.3697***	−0.3861***	−0.3444***	−0.3758***	−0.3413***	−0.2727***	−0.2442*	−0.3220**
	(0.1327)	(0.1667)	(0.0908)	(0.1327)	(0.0389)	(0.1081)	(0.0712)	(0.0789)	(0.1486)	(0.1552)
divorced	−0.2907***	−0.3176***	−0.4156***	−0.4460***	−0.3444***	−0.3833***	−0.2779***	−0.2331***	−0.1395***	−0.1351**
	(0.0626)	(0.0888)	(0.0521)	(0.0514)	(0.0389)	(0.0506)	(0.0477)	(0.0383)	(0.0573)	(0.0642)
widowed	−0.3450**	−0.2901***	−0.3742***	−0.3666***	−0.3695***	−0.3606***	−0.2281***	−0.1754***	−0.1513***	−0.1406***
	(0.138)	(0.0986)	(0.0583)	(0.0814)	(0.0571)	(0.0552)	(0.0641)	(0.0661)	(0.043)	(0.0414)
never married	−0.2460***	−0.2435***	−0.2444***	−0.2930***	−0.2161***	−0.2501***	−0.1529***	−0.1136***	−0.1343***	−0.1053**
	(0.0604)	(0.0724)	(0.0443)	(0.0456)	(0.0336)	(0.0379)	(0.032)	(0.0284)	(0.0496)	(0.0416)
in a civil partnership	−0.0888	−0.0417	−0.0815	−0.1702	−0.0354	0.0834	0.124	−0.0192	−0.0572	−0.1502
	(0.169)	(0.2588)	(0.148)	(0.1581)	(0.1487)	(0.1506)	(0.1276)	(0.0597)	(0.1198)	(0.1779)
Employment status dummies (reference group: self-employed)										
employed	−0.119	−0.0841	−0.097**	−0.0758	−0.0815**	−0.0515	0.0029	−0.0137	0.0638	0.0333
	(0.0904)	(0.0869)	(0.0399)	(0.0482)	(0.0424)	(0.0395)	(0.0451)	(0.0288)	(0.0645)	(0.0440)
unemployed	−0.9504***	−0.9764***	−0.7366***	−0.8859***	−0.5092***	−0.5555***	−0.2996***	−0.3590***	0.0437	0.0998
	(0.1802)	(0.2046)	(0.0536)	(0.0926)	(0.0777)	(0.1075)	(0.1155)	(0.0764)	(0.0884)	(0.1209)
retired	−0.2025**	−0.0707	−0.0975*	−0.1121	0.0429	0.0595	0.1352***	0.1818***	0.1732***	0.1084*
	(0.0899)	(0.1123)	(0.0523)	(0.0865)	(0.0722)	(0.0538)	(0.0509)	(0.0544)	(0.054)	(0.0601)
maternity leave	0.8698**	0.7691**	0.8098***	0.8348***	0.628***	0.6015***	0.5596***	0.4157***	0.6373***	0.4740**
	(0.3712)	(0.3389)	(0.1992)	(0.1213)	(0.1231)	(0.1455)	(0.2031)	(0.1161)	(0.1897)	(0.2121)
family care	−0.2891**	−0.2998**	−0.2359***	−0.2227***	−0.1319***	−0.1521**	−0.0498	−0.0242	0.0726	0.0952
	(0.1236)	(0.1299)	(0.0877)	(0.0748)	(0.08)	(0.0671)	(0.0755)	(0.0561)	(0.0995)	(0.0658)
sick, disabled	−1.1895***	−1.2743***	−0.8602***	−1.0215***	−0.6748***	−0.9488***	−0.6368***	−0.7928***	−0.3377***	−0.3089*
	(0.1842)	(0.1546)	(0.0677)	(0.0994)	(0.0579)	(0.0867)	(0.073)	(0.1054)	(0.1012)	(0.1623)
Pseudo-R squared	0.144	0.1833	0.1309	0.1543	0.1089	0.1191	0.1179	0.0308	0.0899	0.09409
Observations	13,348	13,348	13,348	13,348	13,348	13,348	13,348	13,348	13,348	13,348

Note: Parameters are given for selected variables. TORQUE standard errors are obtained with delta method, quantile regression standard errors are bootstrapped.
*p<0.1; **p<0.05; ***p<0.01

happiness. Contrary to what economic theory assumes, psychologists believe income and other life circumstances have negligible impact on happiness, and that happiness is determined by genetics. I show that both of these theories are partially correct. Results of this paper indicate income has statistically significant effect on happiness for most part of the conditional distribution, but the amount of happiness bought by extra income diminishes with quantiles of happiness. Further, I show Easterlin was right all along: money does not buy happiness. But this is true only for the happiest people. Diminishing marginal effects property holds for other variables as well, possibly implying that external factors are not very important for the happiest people because they are happy just by their nature. This finding partly confirms psychologists' "set-point" theory.

TORQUE estimators are compared with ordinary quantile regression estimators. The estimated link function shows that the distance between happiness levels is not equal which means that the continuity assumption of the dependent variable does not hold in this application. The largest disparities between $\hat{\beta}_{\tau}^{TORQUE}$ and $\hat{\beta}_{\tau}^{QR}$ occur at the extremes of the distribution, with overestimation at lower tails and underestimation upper tails. The estimated difference between quantile regression estimator and TORQUE estimator is 4-74 percent for $\tau < 0.5$ and for 5-53 percent for $\tau > 0.5$! Note that the consistency of the results relies only on the endogeneity assumption and that the model is linear in parameters. I also show that TORQUE outperforms OLS and standard quantile regression and the gain is substantial.s

A number of challenges remain for a future research. First, it is a well known fact that well being is affected by relativities (A.Clark et.al 1996, 2008; A.Ferrer-i-Carbonell et.al 2014). For example, additional income may not increase happiness if the incomes of the comparison group also increase. Easterlin (1995) notes, that if relative income dominates absolute income effects, this could explain why cross-sectional analysis predicts that richer people are happier. Therefore, the inclusion of relative income as one of the predictors can achieve better identification. Next, as with any cross-sectional data, it is very challenging to make a strong causal claim.

Another potentially useful modification would be the analysis within longitudinal dataframe. An application of fixed effects TORQUE could potentially better deal with endogeneity issue by exploiting the panel properties of a data set and capture time invariant unobserved individual effects. Finally, Hong and Zhou (2013) developed a multi-index model for quantile regression which accounts for correlation between covariates and residuals. The application of such new and more robust statistical methodologies in the same framework could complement results of this paper.

Chapter 2

Fertility in the Wake of Better Schools: an Application of Transformed Ordinal Quantile Regression

2.1 Introduction

The analysis of fertility along both intensive (number of children a woman has) and extensive[1] (the decision to become a mother) margins is very important for policy evaluations such as family planning and returns to education since exogenous shocks to fertility do not always cause both margins to respond in the same direction. For example, Becker and Lewis's (1973) standard quantity-quality model of fertility predicts that a decrease, say, in children's health care costs leads to a woman choosing a fewer number of children because of substitution effect. But this is true only along the intensive margin. At the extensive margin fertility should increase because lower health care costs boost the opportunity to invest in child quality. Consequently, more

[1]Extensive margin, motherhood, childlessness are used interchangeably throughout the paper

childless women decide to become mothers. These types of shocks partially cancel each other out, therefore, a decomposition of fertility into extensive and intensive margins helps to explain cross-sample differences in average fertility.

In this paper I study the changes in fertility along both intensive and extensive margins in response to improved schooling opportunities. Better schooling circumstances represent an increase in the returns to investment in child human capital. This work is closely related to the paper of Aaronson, Lange &Mazumder (2014) (AL&M). Similarly, I focus on augmented quantity-quality approach by allowing for the change in fertility at the extensive margin for Southern rural black women caused by the Rosenwald Rural School Initiative. The model predicts that fertility increases along the extensive margin as the opportunity to invest in child quality expands. AL&M define "extensive" margin as the probability that a woman has a child, whereas "intensive" margin is defined as the number of children a woman has, conditional on having at least one child. They divide their sample according to these definitions and use diff-n-diff OLS estimation. By directly applying OLS, AL&M implicitly assume that fertility is a continuous variable and that the decision of having the second child is equivalent to the decision of having, for example, the fifth child. This paper differs from AL&M by not imposing that fertility is a continuous variable. Further, to deal with possible endogenous school selection, AL&M control with rich set of covariates, including country-fixed effects and time trends. When some of the covariates are highly correlated it may result in high variability in the estimates of the regression coefficients. To address this potential multicollinearity problem, I consider l_1-penalized quantile regression (l_1-QR) which uses the sum of absolute values of coefficients as penalty. l_1-QR has the advantage of simultaneously controlling the variance of the fitted coefficients while performing variable selection. Then I apply TORQUE to the model selected by l_1-QR (post-l_1-QR). TORQUE relaxes the assumption that the number of children is a continuous variable. TORQUE technique generalizes ordinary quantile regression providing richer description of the data and allowing for the analysis of marginal effects of explanatory variables along different

quantiles of the conditional distribution of fertility. This property enables the study fertility changes along both intensive and extensive margins without sample selection.

Fertility is an ordinal response variable. Ordered probits and logits handle this type of data well. However, as any Maximum Likelihood Estimators these methods can be sensitive to distributional misspecification. TORQUE estimators are more robust than Maximum Likelihood Estimators because TORQUE methodology does not rely on parametric assumptions of the conditional distribution function, thus reducing model misspecification biases.

Most of the fertility studies can be divided into two main streams.[2] The first stream examines only the intensive margin or the average birthrate. If the relationship between an explanatory variable and fertility moves in opposite directions for two margins, then focusing only on average fertility can lead to underestimation because these two effects can partially or completely cancel each other out. For example, C.Avitable et.al (2012) find that the introduction of birthright citizenship rights in Germany ultimately lead to a reduction in immigrant average fertility. New migration policy represented a positive shock to the returns to investment in child human capital which can be translated into reduction in the cost of child's "quality". According to Becker's quantity-quality model, the policy should decrease fertility rates. However, average fertility rates cannot capture policy effects at the intensive and extensive margin because lower child quality cost incentivizes childless women to become mothers. M.Bailey et.al (2011) examine the hypothesis that advances in household technology caused the US baby boom and find no support for this claim. Hansen et.al (2014) investigate effects of education on fertility transitions in the US using System GMM. Both of those papers use average fertility measures which cannot

[2]G.Becker et.al (1973, 1990), J.Hotz et.al(1988), L.Whittington et.al (1990), C.Avitable et.al (2012), La Ferrara et.al (2012), M.Shields et.al (1986), G.Boyer (1989), M.Bailey et.al (2011), S.Hansen et.al (2014) study the fertility without distinguishing intensive and extensive margins, R.Baughman et.al (2003) P.Gobbi (2013), C.Hakim (2013) study theoretical and empirical evidence of childlessness and motherhood.

capture the effect of exogenous shocks on the decision of motherhood possibly leading to an underemphasis of that shock on fertility.

The second strand of the research focuses on only the extensive margin. Childlessness is key to understanding differences in average fertility across countries, over time, or across education groups (Baudin et.al, 2015). The fertility literature distinguishes between voluntary and involuntary childlessness. Voluntary childlessness happens when a woman can biologically and financially afford to have children but decides not to. Involuntary childlessness occurs due to biological constraints or poverty. Baudin et.al (2015b) develop a structural model to identify involuntary childlessness. They study the extensive margin of fertility and the effectiveness of development policies. Patterns between fertility and marriage, fertility and education, and education and marriage are used to identify the parameters of the theory. Baughman et.al (2003) study the effect of earned income tax credit on motherhood. Hakim (2003) develops a theory that explains the increase of childlessness with more career oriented women, giving a priority over family life. For practical reasons most studies assume that childlessness is a choice. Similarly to these papers, I assume motherhood is a choice. This assumption is not very restrictive given a setting is within richer countries. As was noted by Gobbi (2013) and Hakim (2003), voluntary childlessness drives the dynamics of motherhood in developed countries.

The economic analysis of fertility is beginning to shift to studies that consider both margins. Baudin et.al (2015a, 2015b) develop a theory of fertility distinguishing between intensive and extensive margins. They fit their model to identify effects of education and marriage on fertility and childlessness in a unified framework. Baudin et.al (2015b) finds that neglecting the endogenous response of marriage and extensive margin of fertility leads to overestimation of the effectiveness of family planning policies. Gobbi (2013) shows higher instances of childlessness and lower fertility following a reduction in the gender wage gap and an increase in the fixed cost of becoming a parent in developed countries. Momota (2015) investigates the relationship between fertility, capital accumulation and economic welfare. All these

studies illustrate the importance of separating the effect on birth decisions into two margins. Otherwise, interpretations may be confounded, particularly when the variable of interest generates opposing effects on the decision to become a parent compared to the decision to have additional children.

To examine effects of Rosenwald schools on fertility I use two samples of women. The first sample includes women who were too old to attend Rosenwald schools themselves but their children had the opportunity to study in these schools. For this cohort of women, school construction decreased the cost of a child quality. AL&M's augmented quantity-quality model predicts that in this case fertility declines at the intensive margin but increases along the extensive margin. Intuitively, higher opportunity to invest in a child quality causes a substitution effect at the intensive margin: women substitute quantity for quality. At the same time, more childless women decide to become mothers because it is necessary to have at least one child in order to invest into his quality. For the second sample I include younger women who themselves had the opportunity to attend Rosenwald schools and get a better education. The construction of schools improved labor market conditions for the younger cohort. In this case, the model predicts that fertility should decline along both margins because of an increase in opportunity cost of having children.

Results of this study roughly confirm the augmented quantity-quality model. At the extensive margin, a better access to education increased the older cohort's probability of having a child by 3.3-4.2 percent. TORQUE results suggest Rosenwald schools did not have statistically significant effect on fertility decisions of a woman who had more than one child. However, OLS estimates infer large and statistically significant negative effects. OLS estimates are larger than that of TORQUE's by 17 to 75 percent. I also test the same theoretical model on a sample of women who could have attended Rosenwald schools themselves. I find that the exposure to a better education significantly decreased the probability of having a child. Results are consistent along both intensive and extensive margins confirming predictions of the model.

The rest of the paper is organized as follows. Section 2.2 provides a simple theory of augmented quantity-quality model of fertility. Section 2.3 describes data and econometric specification. Section 2.4 presents results. Section 2.5 concludes and discusses implications of the results.

2.2 Theory

Assume that households derive their utility from consumption c, number of children n, and the quality of those children q.[3] Each household is endowed with one unit of time. The utility is maximized subject to a time constraint which can be devoted to earn income y and raise children:

$$\max_{c,n,q} U(c,n,q) \qquad s.t \qquad c + yn(\tau^q q + \tau^n) \leq y.$$

τ^q is the fraction of household time endowment required for each unit of a child's quality. The quality can be thought of as education, better nutrition, extracurricular activities, etc. τ^n is the time fraction that is spend to raise a child, which is fixed and independent of child's quality. The opportunity cost of raising a child, or equivalently, child's total price is therefore equal to $y(\tau^q q + \tau^n)$.

The optimal number of children is derived from:

$$U_n = MC_n = P_n = y(\tau^q q^* + \tau^n)$$

$$U_q = MC_q = P_q = yn^*\tau^q$$

Here, U_n is the marginal utility of the next child, MC_n is the marginal cost of having an additional child. U_q is the marginal utility from increasing the quality of each child in the family by one unit, and MC_q is the marginal cost of the quality. Marginal costs are also referred to as shadow prices of quality and number of children, P_q and P_n,

[3] Theoretical framework is based on AL&M and Galor (2012)

respectively. It can be seen that a boost in household income raises shadow prices of quantity and quality. Equivalently, a positive income shock increases the opportunity cost of raising a child to which women respond with reducing their fertility.

On the other side, an increase in household investments in quality raises the shadow price of quantity. This effect generates Becker and Lewis's (1973) quantity-quality substitution: larger spending on quality tends to reduce the optimal number of children. For practical convenience it is very common to introduce an assumption that fertility is always positive:

$$\lim_{n^+ \to 0} U_n = \infty$$

This assumption leads an analysis to focus exclusively on fertility along the intensive margin. However, the interaction between quality and quantity leads to Beckers quantity-quality trade-off. This trade-off occurs at high and low fertility levels. This implies that around the extensive margin, quantity and quality must be complements since having a child is a prerequisite to investing in child quality. AL&M refer to this complementarity at the extensive margin as the "essential" complementarity. Note, the value of being childless is only a function of income, $V_0(Y)$, but the value of having children depends positively on income and negatively on the time costs of quality and quantity, $V(Y, \tau^q, \tau^n)$. Women choose to have children if $V \geq V_0$. If τ^q declines, then more women will choose to become mothers, so fertility increases along the extensive margin. Because quantity and quality are substitutes at the intensive margin, women will choose to have less children overall.

2.3 Data and Econometric Specification

2.3.1 Dependent variable

To test fertility changes along both margins I combine the information on fertility with the information on the availability of schools. The dependent variable is the

total fertility constructed as the number of surviving children under the age of ten.[4] The reason for the age limitation is to avoid an issue associated with children leaving their home.

Fertility of women who attended schools themselves were affected by the school exposure differently than women who did not attend but whose children had the opportunity to attend school. Rosenwald schools improved labor market conditions for women able to attend (i.e the younger cohort), increasing the opportunity cost of having children. On the other hand, exposure to schools decreased the quality cost of raising a child for women who did not attend the school themselves. To separate these two effects I use two distinct samples of women: older cohort and a younger cohort. The data for the older cohorts is taken from 1.4 percent sample for 1910, 1 percent sample for 1920, and 5 percent sample for 1930 censuses using the Integrated Public Use Microdata Series (IPUMS) and includes only women who were 25-49 of age at the time of the census. The sample for the younger cohort is drawn from IPUMS 1930 5 percent sample and consists of women who were born between 1908 and 1912. I did not include later years to avoid potential sample selection bias caused by "Great Migration" of blacks from South to North. The issue with only using 1930 census data for the younger cohort is the inability to observe fertility decisions beyond 1930. If the exposure to schools just delayed childbirth for this cohort, then my estimates are not capturing this effect.

In contrast to AL&M, I do not construct separate samples to capture exposure effects on fertility along intensive or extensive margins. The methodology I use in this paper allows the estimation of the effects of Rosenwald schools along different quantiles of the conditional distribution of fertility without the sample selection.

[4]See AL&M for detailed data and variable selection strategies.

2.3.2 Rosenwald schools and school exposure

The second source of information describes the exposure to Rosenwald schools. Between 1912-1932, nearly 5000 Rosenwald schools were built for the black children of the rural South. They were built in Oklahoma, Kentucky, Missouri, Maryland, as well as in the eleven states of the Confederacy. The schools were named after Julius Rosenwald, a prominent humanitarian and source of funding. Equally important in Rosenwald Initiative was Booker T. Washington, who was the first principal of the Tuskegee Normal School for Colored Teachers in Alabama. Washington convinced Rosenwald to provide funding for school construction for rural black children. Washington had contacts and knowledge, Rosenwald had money.

The schoolhouses were designed by Tuskegee architects. At the time these buildings were equipped with modern amenities such as adequate sanitation, ventilation, and windows large enough for reading light. When the funding ended in 1932, it had served more than 660,000 students, according to National Trust for Historic Preservation. The capacity was large enough to accommodate more than one-third of the child population of the rural South.

Identification relies on the variation in the exposure to Rosenwald schools combined with variation in fertility. Women in the sample are connected to Rosenwald schools through their county of residence, birth year and an urban/rural status. The coverage of the schools $C_{c,t}$ is measured by the ratio of the product of Rosenwald Fund's count of Rosenwald teachers in a given county c for the year of t and the average class size of forty-five relative to the number of black children between the ages of seven and thirteen in that county c and a year t. Then the exposure is defined as the average coverage for each cohort of women. For older cohort the exposure is equal to $E_{c,t} = \frac{1}{10}\sum_{n=1}^{10} C_{t-n,c}$, and for younger cohorts the exposure is $E_{c,b} = \frac{1}{7}\sum_{n=1}^{7} C_{b+n+6,c}$.

Summary statistics for exposure and other measures are given in Table 2.1. The average age of women in the older group is 35.5 years old and 20 years old in the

Table 2.1: Summary statistics

	Older Cohort			Younger Cohort		
Variable Name	All	Black-Rural	Black-Urban	All	Black-Rural	Black-Urban
Total Fertility	1.11	1.23	0.49			
	(1.43)	(1.65)	(1.07)			
Total Fertility (1910)	1.38	1.51	0.61			
	(1.56)	(1.73)	(1.13)			
Total Fertility (1920)	1.22	1.34	0.51			
	(1.47)	(1.66)	(1.05)			
Total Fertility (1930)	1.02	1.13	0.47	0.33	0.4	0.21
	(1.38)	(1.62)	(1.07)	(.72)	(.85)	(.61)
Rosenwald exposure	.14	.13	.19			
	(.21)	(.17)	(.24)			
Rosenwald exposure (1910)	0	0	0			
Rosenwald exposure (1920)	.00047	.0058	.0047			
	(.15)	(.15)	(.15)			
Rosenwald exposure (1930)	.2008	.19	.2346			
	(.23)	(.18)	(.25)			
Own exposure				.071	.074	.076
				(.12)	(.1)	(.12)
Age	35.46	35.38	34.96	19.94	19.88	20
	(7.08)	(7.13)	(6.97)	(1.43)	(1.43)	(1.42)
Age married	20.68	20	20.56	17.59	17.46	17.45
	(4.6)	(1.65)	(5.19)	(1.82)	(1.83)	(1.88)
Black	.26	.17	.1	.29	.2	.1
Observations	382,417	63,740	38,929	111,711	22,279	10,588

Note: Sample for older cohort was constructed from IPUMS 1910, 1920, and 1930 year and includes 25-49 years old women. Younger cohort sample was constructed using IPUMS 1930 and includes 18-22 years old women. For more details on how variables are constructed refer to the text or AL&M.

younger cohort. Interestingly, women in the younger group typically married at 17.5 years of age, three years earlier than the previous generation. Blacks comprise only 26-29 percent of the total sample. Over the period of 1910-1930 fertility for all subgroups declined by 23-26 percent. Also, the younger women had one less child on average, but since the sample period ends while they are still able to bare children it is possible they had as many or more children than the older cohort. Table shows a rapid increase in Rosenwald school exposure for the older cohort, rising from 0 percent in 1910 to 23.5 percent in 1930. Exposure for rural black women in the young cohort averages 7.4 percent.

To build schools Rosenwald did not simply give money: he required people to show how much they wanted a school. People had to raise additional cash, contribute labor and convince local white government to commit public funds. This funding structure suggests that location of schools was likely not random. However, Aaronson and Mazumder (2011) show that black socioeconomic characteristics do not predict the location of the schools and that selection bias is small. Further, to address the non random selection issue AL&M control for a rich set of covariates, including time trends, county fixed effects, and interactions of exposure with race and rural status. Interaction terms exploit the explicit targeting of the program to rural blacks while controlling for urban blacks and rural whites. However, if covariates are highly correlated, this may result in inconsistent estimators. I address this issue by using least absolute shrinkage and selection operator (LASSO) which performs variable selection in order to enhance estimation accuracy. More details on the application of LASSO are given in econometric specification section.

Note, the theoretical and empirical model does not allow for involuntary childlessness. Childlessness can be the result of poverty and the poorest families are more likely to be affected by sub fecundity factors. As a result, for some of the poorest younger cohort the exposure to schools did not cause their fertility to decline. Conversely, improved labor market conditions provided sufficient income to

procreate. Ignoring involuntary childlessness in the given framework underestimates the relationship between school exposure and fertility.

2.3.3 Econometric specification

Explanatory variables explain the variation in fertility of a woman i in county c in census time t for a given quantile τ of the dependent variable's conditional distribution. The basic empirical specification is:

$$\begin{aligned} y_{ict}^{\tau} = \beta_0^{\tau} + \beta black_i^{\tau} + \beta_2^{\tau} rural_i + \beta_3^{\tau} blackrural + \beta_4^{\tau} X_i + \beta_5^{\tau} age_{it} + t + c \\ +(\gamma_0^{\tau} + \gamma_1^{\tau} black_i + \gamma_2^{\tau} rural_i + \gamma_3^{\tau}(black_i * rural_i)) * E_{tc} + e_{ict}^{\tau} \end{aligned} \tag{2.1}$$

The measure of exposure to Rosenwald schools is interacted with race and rural status dummies. I use this specification to estimate four different estimators of Rosenwald schools exposure effect on fertility. The first, $\hat{\gamma_0} + \hat{\gamma_1}$ provides diff-n-diff estimator of the schools' effect on black women's fertility. To difference out effects that possibly impacted only blacks and were unrelated to school introduction, I use $\hat{\gamma_1} + \hat{\gamma_3}$. Third, $\hat{\gamma_2} + \hat{\gamma_3}$, uses the difference between rural blacks and rural whites in order to remove any common rural effect. Finally, $\hat{\gamma_1} + \hat{\gamma_2} + \hat{\gamma_3}$, is the triple difference estimator which controls for common rural and race effects. This is the preferred estimator. Analogous specification is also constructed for the younger cohort.

In this paper I use a generalized version of quantile regression. Quantile regression is an important statistical method for analyzing the impact of regressors on the conditional distribution of the dependent variable. It captures the heterogeneous impact of covariates, exhibits robustness to outliers, and has excellent computational properties. Furthermore, in this paper, quantile regression allows an estimate of the effects of the exposure to school on the fertility along both margins without sample selection. However, quantile regression relies on the assumption that the conditional distribution of the response is continuous. The main problem with the estimation of conditional quantiles of fertility is that dependent variable has a discrete

distribution so $Q_y(\alpha|x)$ cannot be a continuous function of the parameters of interest. TORQUE[5] overcomes this limitation by constructing a continuous random variable whose quantiles have one-to-one relation with the quantiles of the dependent variable. A variable satisfying this requirement is constructed by jittering the response variable.

Note, this empirical specification contains a very rich set of covariates, including over 1000 county fixed effects. When some of the covariates are highly correlated it may result in high variability in the estimates of the regression coefficients. To address potential multicollinearity problem, I consider LASSO quantile regression. LASSO quantile (l_1-QR) is a regression analysis method that performs variable selection by penalizing l_1-norm of parameter coefficients in order to enhance prediction accuracy. LASSO shrinks some of the coefficients while forcing others to be zero, effectively choosing a simple model that does not include variables with zero coefficients. However, l_1-QR's penalized estimator has a regularization bias. To deal with potential bias problem, I use post-penalized estimator (post-l_1-QR) which applies ordinary, unpenalized quantile regression to the model selected by l_1-QR. Estimates of l_1-QR can be calculated by solving:

$$\arg\min_{\beta} \sum_{i=1}^{n} \rho_\tau(y_i - \beta_0 - X_i^T\beta) + \lambda||\beta|| \tag{2.2}$$

The first part of the equation is exactly Koenker and Bassets criterion function[6]. The second part penalizes criterion function using l_1 norm of β. $\lambda > 0$ is the regularization paprameter that balances the quantile loss and the penalty. Larger λ values impose more penalty and less covariates are chosen. Equivalently, if $\lambda = 0$, we obtain no shrinkage. Instead of exogenously choosing λ I use data driven choice of the penalty level. This method is proposed by A. Belloni et.al (2011) and works as follows:

[5]For a detailed explanation of TORQUE and its estimation please refer to Section 1.2

[6]As defined by equation 1.2

$$M = n \sup_{\tau} \max_{1 \geq j \geq p} \left| E_n \left[\frac{(x_{ij}(\tau - 1u_i \leq \tau)}{\hat{\sigma}_j \sqrt{\tau(1-\tau)}} \right] \right|, \tag{2.3}$$

where p is the total number of regressors, $\hat{\sigma}_j = \sqrt{E_n(x_{ij}^2)}$, and $u_1, ..., u_n$ are i.i.d. uniform (0,1) random variables, independently distributed form x. The random variable M has a known distribution conditional on X. The penalty λ is set to

$$\lambda = c * M(1 - \alpha | X) \tag{2.4}$$

where $M(1-\alpha|X) := (1-\alpha)$-quantile of M conditional on X, and the constant $c > 1$ depends on the design. $M(1 - \alpha|X)$ is computed using simulation, and $1 - \alpha$ is the confidence interval which can be set to be equal to 0.1. The formal rational for using (2.3) for the penalty level is that this choice leads to the optimal rates of convergence of l_1-QR.

TORQUE is applied to the model selected by l_1-QR (post-l_1-QR).

2.4 Results

2.4.1 The Analysis of the Link Function Λ

Figure 2.1 shows the plot of the estimated function of $\Lambda(\tilde{y})$ for older cohort.

From the figure it is obvious that the association between $\hat{\Lambda}$ and $\tilde{Y}$ resembles a piecewise linear relationship with the discontinuity at $\tilde{y} = 1$. The slope of $\hat{\Lambda}(\tilde{y})$ is larger at $\tilde{y} < 1$ relative to $\tilde{y} > 1$. Steeper slope suggests that the distance along the extensive margin is larger relative to the intensive margin and that coefficients along the extensive margin are most likely to be underestimated if we assume data continuity. The relationship between $\hat{\Lambda}$ and $\tilde{Y}$ reveals that the decision to have one more child at higher levels of fertility is relatively easier compared to the decision

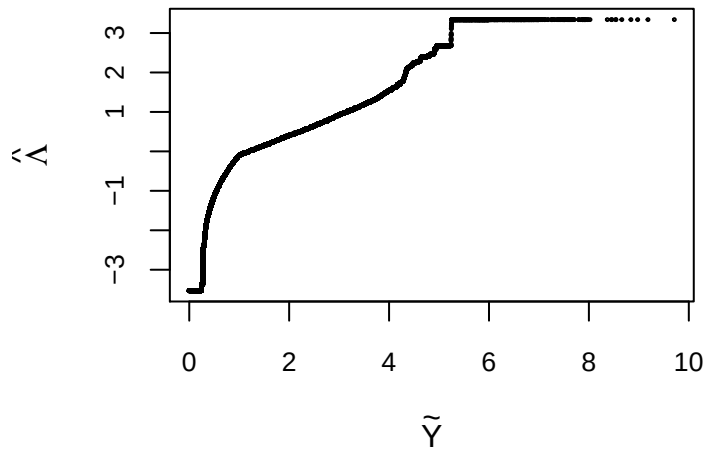

Figure 2.1: Estimation results of Lambda for older cohort

facing a childless women. In other words, it takes more to become a mother for the first time. [7]

Estimated function of $\Lambda(\tilde{y})$ (Figure 2.2) for younger cohorts tells similar story.

2.4.2 Main Results

TORQUE methodology estimates of the effect of Rosenwald exposure on overall fertility for older women are presented in Table 2.2. Note, TORQUE estimators are recovered in the following way. First, I regress $\hat{\Lambda}_i$ on $\tilde{Y}_i$: $\hat{\Lambda}_i = a + c\tilde{Y}_i$ for $\tilde{Y}_i < 1$. Next, I rearrange to get: $\tilde{Y} = \frac{\hat{\Lambda}}{\hat{c}} - \frac{\hat{a}}{\hat{c}}$. Finally, $\frac{\delta Y_\tau}{\delta x} = \frac{\hat{\beta}_\tau}{\hat{c}}$. Fitted values and $\hat{\beta}_\tau$ are obtained using (1.3) of the model. I repeat the same steps for $\tilde{Y}_i > 1$ as well.

About 58 percent of the sample do not have any children, 16 percent of women have only one child, and about eigth percent of women have three kids. To estimate

[7] Because the portion of observations for older cohorts with number of kids larger or equal to five is only 0.4 percent, quantile estimation is not reliable at 99.6 percentile or higher. Similarly, for younger cohort Λ flats out around $\tilde{y} = 4$; the portion of observations with $\tilde{y} > 4$ is only 0.6 percent. In my further estimations I drop these observations.

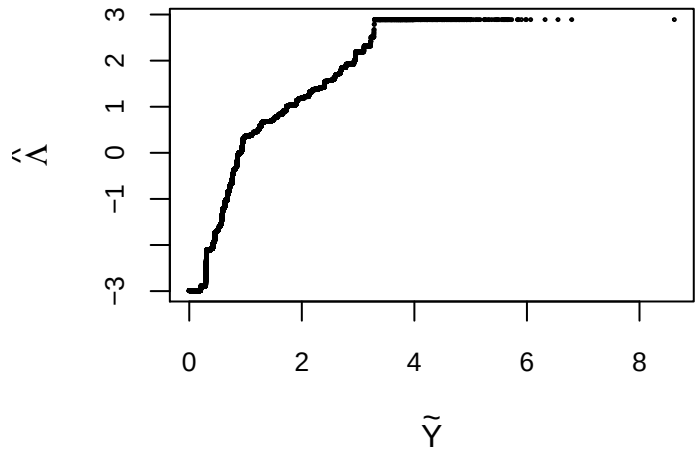

Figure 2.2: Estimation results of Lambda for younger cohort

the school exposure effect around the extensive margin, I set the quantile to 60. To estimate exposure effects for women with one, two, and three children I set quantiles to 68, 85, and 95, respectively.

The main variable of interest is triple differenced estimate γ_3. As in AL&M, it indicates the change in probability of having a child going from no exposure to complete exposure for rural black women. The sample for our older cohort consists of women who did not attend the school themselves, but their children had the opportunity for a better quality education. School exposure lowered the cost of child quality. In that case, augmented quality-quantity model predicts that fertility decreases along the intensive margin because women substitute the quantity of their children for the quality. According to model predictions γ_3 should be negative at the intensive margin. Meanwhile, at the extensive margin model predicts an increase in fertility as women take advantage of the reduced cost of child quality. In this case, γ_3 should be positive.

Table 2.2: TORQUE model results of older cohort's predicted fertility

	Dependent variable: Total Fertility			
	Extensive Margin	Intensive margin		
	(0)	(1)	(2)	(3)
	Triple differenced estimator			
Rosenwald exposure (γ_0)	$-.014^{**}$	$-.094^{***}$	$-.128^{***}$	$-.142^{***}$
	(.006)	(.031)	(.039)	(.054)
exposure_black (γ_1)	.001	.006	$.111^{*}$	$-.051$
	(.011)	(.056)	(.069)	(.094)
exposure_rural (γ_2)	.008	$.079^{**}$	$.1^{**}$	.096
	(.008)	(.04)	(.05)	(.069)
exposure_blackrural (γ_3)	$.033^{**}$	$.16^{**}$	$-.104$	$-.056$
	(.016)	(.083)	(.102)	(.138)
	Diff-n-diff estimator			
A. Exposure, black				
Rosenwald exposure (γ_0)	$-.007^{*}$	$-.038^{*}$	$-.056^{**}$	$-.08^{**}$
	(.004)	(.021)	(.026)	(.036)
exposure_black (γ_1)	.009	.042	.051	$-.088$
	(.008)	(.041)	(.05)	(.067)
B. Exposure black, Rural-Urban				
exposure_black (γ_1)	$-.011$	$-.064$	.006	$-.174^{**}$
	(.01)	(.049)	(.057)	(.085)
exposure_blackrural (γ_3)	$.042^{***}$	$.216^{***}$	$-.01$	.35
	(.014)	(.074)	(.086)	(.124)
C. Exposure rural, Black-White				
exposure_rural (γ_2)	$-.005$	$-.004$	$-.019$	$-.28$
	(.005)	(.027)	(.034)	(.047)
exposure_blackrural (γ_3)	$.035^{***}$	$.16^{***}$	.007	$-.103$
	(.012)	(.06)	(.075)	(.1)
Observations	309,418	309,418	309,418	309,418
Pseudo R^2	.073	.077	.08	.098

*p<0.1; **p<0.05; ***p<0.01

Note: Pseudo R^2 are reported for triple differenced estimation.

Results form TORQUE for older cohort are given in Table 2.2. Column one of the table shows the effects of Rosenwald exposure on fertility along the extensive margin in the last ten years. For triple differenced specification our preferred estimator indicates that complete exposure to schools increases the probability that a woman had a child in the preceding ten years by 3.3 percentage points which is statistically significant at ten percent level. The effects for alternative estimators (γ_3 for B and C rows) are 4.2 and 3.5 percent respectively and statistically significant at one percent level. Column two of the Table 2.2 reports estimates for woman who had only one child. Exposure to schools increased the probability of having another child from sixteen to almost twenty-two percentage points. Empirical results for women with more children weakly confirm model predictions: most of the effects have a negative sign but none of them are statistically significant from zero. Rosenwald school exposure (γ_0) had significantly negative effects on fertility for all woman while controlling for other factors.

Table 2.3 shows results from an OLS regression. Note, I apply OLS to the model selected by LASSO. Column one reports the effect of Rosenwald exposure on overall fertility in the last ten years. The preferred triple difference estimator indicates exposure increased the probability of having a child from 5.4 to 10.8 percentage points. However, these estimates are not always statistically significant. Column two reports the effects at the extensive margin. Recall, the response at the extensive margin is an indicator of having at least one child by age nine. γ_3 ranges from 4.9 to 5.8 and it is statistically very significant in all of the specifications. Column three reports results along the intensive margin. The response for this subsample indicates the number of total children conditional of having at least one child before exposure. γ_3 has a negative sign in all cases which confirms the prediction of the model.

Now I compare TORQUE results with OLS results. From the figure 2.1 we should expect that at the extensive margin OLS estimates are larger than TORQUE estimates and smaller at the intensive margin. A comparison of the estimates confirms the prediction. At the extensive margin OLS estimates are larger than TORQUE

Table 2.3: OLS model results of older cohort's predicted fertiltiy

	Dependent variable:		
	Total Fertility	extensive	intensive
	(1)	(2)	(3)
	Triple differenced estimator		
Rosenwald exposure (γ_0)	$-.089^{***}$	$-.015^{**}$	$-.122^{***}$
	(.021)	(.007)	(.029)
exposure_black (γ_1)	$-.01$	$-.009$	.047
	(.036)	(.013)	(.065)
exposure_rural (γ_2)	$.058^{**}$	$-.007$	$.13^{***}$
	(.027)	(.01)	(.035)
exposure_blackrural (γ_3)	.063	$.058^{***}$	$-.158^{**}$
	(.054)	(.019)	(.084)
	Diff-n-diff estimator		
A. Exposure, black			
Rosenwald exposure (γ_0)	$-.047^{***}$	$-.016^{***}$	$-.03^{*}$
	(.014)	(.005)	(.018)
exposure_black (γ_1)	.006	.014	$-.055$
	(.027)	(.01)	(.41)
B. Exposure black, Rural-Urban			
exposure_black (γ_1)	$-.078^{**}$	$-.02^{*}$	$-.051$
	(.032)	(.011)	(.061)
exposure_blackrural (γ_3)	$.108^{**}$	$.049^{***}$	$-.048$
	(.049)	(.017)	(.078)
C. Exposure rural, Black-White			
exposure_rural (γ_2)	$-.024$	$-.021^{***}$	.015
	(.019)	(.007)	(.021)
exposure_blackrural (γ_3)	.054	$.049^{***}$	$-.11^{**}$
	(.041)	(.014)	(.053)
Observations	373,044	373.044	182.067
R^2	.125	.11	.1

*p<0.1; **p<0.05; ***p<0.01

Note: R^2 are reported for triple differenced estimation.

estimates by 17 to 75 percent. OLS results at the intensive margin are always smaller than TORQUE estimates for women with more than one child. However, OLS estimates are almost always statistically significant, whereas TORQUE estimates have mixed signs and mostly are not different from zero. We see the evidence broadly suggesting that response along the extensive margin dominated the effect along the intensive margin. This implies that the total number of children growing up in black rural families increased. Further, the OLS result implies that the distribution of number of children was compressed from both ends, while TORQUE estimates suggest the compression from only one tail. Finally, TORQUE estimates suggest that Rosenwald school exposure increased total fertility for Southern rural black woman by more than the OLS results would suggest.

Table 2.4 presents TORQUE results for the younger cohort of women. It reports the effect of Rosenwald school exposure when a woman was at the age of seven to thirteen on her fertility from the ages of 18 to 22. Better access to higher quality education at Rosenwald schools during childhood increased the human capital of women who attended these schools (Aaronson and Mazumder, 2011). Improved labor market conditions most likely increased the opportunity costs of having a child for this cohort. The model predicts, that for a county that goes from zero exposure to a full exposure, Rosenwald school effects on fertility should be negative along both intensive and extensive margins.

Column one of the table 2.4 reports exposure effects on fertility along the extensive margin. The fraction of women with more than two children is only 0.6 percent which makes quantile estimates at very high quantiles unreliable. For effects along the intensive margin I report estimates for woman having one and two children which are shown in columns two and three. γ_3 estimates along the extensive margin are negative and statistically very significant except for the last specification. γ_3 estimate for rural black-rural white is positive, although not statistically significant from zero. Along the extensive margin the exposure to schools increased the probability of becoming a mother by eight to eleven percentage points. Empirical results at the intensive margin

Table 2.4: TORQUE model results of younger cohort's predicted fertility

	Dependent variable: Total Fertility		
	Extensive Margin	Intensive margin	
	(0)	(1)	(2)
	Triple differenced estimator		
Rosenwald exposure (γ_0)	.022*	.0	−.169*
	(.013)	(.06)	(.102)
exposure_black (γ_1)	.092***	.182	.485***
	(.024)	(.114)	(.181)
exposure_rural (γ_2)	−.036**	−.038	.212*
	(.016)	(.027)	(.121)
exposure_blackrural (γ_3)	−.078***	−.114	−.619***
	(.03)	(.114)	(.226)
	Diff-n-diff estimator		
A. Exposure, black			
Rosenwald exposure (γ_0)	−.001	−.033	.041
	(.006)	(.036)	(.062)
exposure_black (γ_1)	.055***	.093	−.074
	(.012)	(.073)	(.118)
B. Exposure black, Rural-Urban			
exposure_black (γ_1)	.113***	.187*	.325**
	(.017)	(.097)	(.145)
exposure_blackrural (γ_3)	−.113***	−.17	−.423**
	(.022)	(.125)	(.185)
C. Exposure rural, Black-White			
exposure_rural (γ_2)	−.017**	−.041	.039
	(.008)	(.045)	(.07)
exposure_blackrural (γ_3)	.012	.07	−.139
	(.018)	(.099)	(.143)
Observations	89,884	89,884	89,884
Pseudo R^2	.052	.059	.05

*p<0.1; **p<0.05; ***p<0.01

Note: Pseudo R^2 are reported for triple differenced estimation.

also confirm model prediction: going from zero to complete exposure decreased the probability of having a child in all specifications. The evidence is stronger and has higher statistical significance for women with two kids. Triple differenced estimator of γ_3 predicts that the exposure to schools decreased the probability of having a child by almost sixty-two percent. However, this large effect is most likely an overestimation. The data is limited to observing the fertility of women between the ages of eighteen to twenty-two, thus these women may have had additional children after the sample ends. In addition, if more career oriented women just delayed their childbirth, then estimates are not capturing this effect.

OLS results for younger cohort of women are reported in Table 2.5. On average, exposure decreased the average fertility of rural black women by nine to almost forty percentage points. OLS estimates also confirm augmented quantity-quality model predictions. The point estimate for triple differenced estimator at the extensive margin is negative but not statistically significant. The negative effect on fertility is especially strong for women having more than one child. Results imply better educated women are less likely to have many children because of larger opportunity costs. Triple differenced estimator γ_3 suggest a decrease in probability by fifty-seven percent. The evidence is also strong when differencing across black and white rural women. History suggests declines in fertility during demographic transitions, such as during the sample period, making large families less common. My results confirm this empirically.

TORQUE and OLS estimates for younger cohort strongly suggest that Rosenwald schools had negative effect fertility for Southern Rural Black women. However, magnitudes of two estimates differ. At the extensive margin the difference of triple differenced γ_3 is about eighteen percent. At the intensive margin OLS γ_3 is smaller by eight percent.

Table 2.5: OLS model results of younger cohort's predicted fertility

	Dependent variable:		
	Total Fertility	extensive	intensive
	(1)	(2)	(3)
	Triple differenced estimator		
Rosenwald exposure (γ_0)	.12***	.021	.189
	(.04)	(.023)	(.132)
exposure_black (γ_1)	.14*	−.092**	.198
	(.073)	(.041)	(.249)
exposure_rural (γ_2)	−.169***	−.057**	−.103
	(.04)	(.027)	(.145)
exposure_blackrural (γ_3)	−.219**	−.064	−.572**
	(.093)	(.052)	(.287)
	Diff-n-diff estimator		
A. Exposure, black			
Rosenwald exposure (γ_0)	.002	−.018	.105*
	(.023)	(.013)	(.058)
exposure_black (γ_1)	.03	.06**	−.204
	(.046)	(.026)	(.128)
B. Exposure black, Rural-Urban			
exposure_black (γ_1)	.259***	.114***	.038*
	(.062)	(.035)	(.21)
exposure_blackrural (γ_3)	−.386***	−.121***	−.676
	(.08)	(.045)	(.248)
C. Exposure rural, Black-White			
exposure_rural (γ_2)	−.053**	−.037**	.082
	(.027)	(.015)	(.063)
exposure_blackrural (γ_3)	−.087	.024	−.382***
	(.059)	(.033)	(.148)
Observations	110,092	110,092	24,619
R^2	.077	.069	.075

*p<0.1; **p<0.05; ***p<0.01

Note: R^2 are reported for triple differenced estimation.

2.5 Conclusions

Using transformed ordinal quantile regression model I test an augmented quantity-quality model of fertility. I focus on the effect of Rosenwald schools on fertility along both intensive and extensive margins for black women of the rural South. The paper is closely related to the paper of AL&M, examining the same theory using the same dataset. However, by directly applying OLS, AL&M implicitly assume fertility to be a continuous variable. This assumption is not supported by my results. The methodology I use relaxes continuity assumption and allows estimation of the effects of exposure without sample selection. Further, to deal with multicollinearity problem, I consider l_1-QR, which has the advantage of controlling the variance of estimates while simultaneously performing variable selection. TORQUE is applied to the model selected by l_1-QR.

Results roughly confirm augmented quantity-quality model of fertility. Estimates are also more consistent compared to AL&M results.

The analysis of fertility along both margins generates an additional test for factors driving demographic transitions. Neglecting distinctly different responses of fertility along two margins can lead to overestimation or underestimation of the effectiveness of family planning policies.

Chapter 3

Medical Marijuana Laws and Heterogeneity in Youth Marijuana Smoking Intensity

3.1 Introduction

The number of U.S. states where marijuana use for medicinal purposes is legal has been growing. As of April, 2016, 25 states, including Washington D.C., legalized marijuana for medicinal purposes. It is expected that this number will increase in the near future as more states are considering passing similar laws. The law allows patients with certain conditions to receive medical marijuana. There is a healthy debate whether the presence of medical marijuana laws (MML) encourage marijuana smoking among adolescents. Some note that MMLs may make marijuana easily accessible and lead to increased illegal marijuana use by teenagers. In addition, MMLs may send a mixed message that use of marijuana is acceptable and is not associated with adverse health effects (Hasin et al. 2015). Others argue that marijuana serves as a gateway drug to other illicit drugs, such as cocaine and heroin. Therefore, all

of these issues make it important to understand the relationship between MMLs and illegal marijuana use among youths.

Our goal in this paper is to estimate the relationship between medical marijuana policies and marijuana consumption among high school students. The findings from the empirical literature on the relationship between MMLs and youth recreational marijuana use is mixed. We contribute to the literature in a number of ways. First, unlike previous studies, we focus on the frequency of marijuana use (intensive margin), rather than participation in marijuana use (extensive margin) by teenagers. Using frequency data allows us to understand the relationship between MMLs and marijuana use for different demographics, such as light vs. heavy marijuana users. Some studies suggest the negative effects of marijuana are especially associated with heavy smokers.

Second, we use state Youth Risk Behavior Survey (YRBS) for 1991-2013. The advantage of using state YRBS is that it provides data that is representative at the state-level, while similar other surveys, such as Monitoring the Future (MTF), are representative of the U.S. at the national level. Another survey, the National Survey on Drug Use and Health (NSDUH), provides state-level estimates starting only in 1999. Consequently, it does not cover several states that passed MMLs prior to 1999 (California, Oregon, and Washington).

Third, we use the transformed ordinal quantile regression (TORQUE) from Hong and He (2010). to estimate the effect of medical marijuana policy on marijuana use among high school students. Unlike ordered probit or logit models, the predicted intervals from TORQUE are more informative. In addition, estimation of TORQUE is robust to model misspecifications as it does not rely on any parametric specification of the conditional distributions.

The paper is organized as follows. Section 3.2 provides background on medical marijuana and discusses the main findings from the literature. Section 3.3 describes the data that is the basis of this study. In Section 3.4 we provide the econometric specification. The results are discussed in Section 3.5. Section 3.6 concludes.

3.2 Background

3.2.1 State MMLs

MMLs allows possession and/or cultivation of marijuana for medicinal purposes. California, Oregon, and Washington were the first to legalize marijuana for medicinal purposes. In April of 2016, Pennsylvania became the 25th state, including the District of Columbia, to legalize medical marijuana (Table 3.1). In states that passed MMLs, patients with certain medical conditions (typically, cancer, glaucoma, HIV/AIDs, multiple sclerosis, epilepsy, seizures, Crohn's disease, wasting syndrome, PTSD) are eligible to receive medical marijuana. However, the exact qualifying conditions for medical marijuana vary from state to state. In addition, MMLs regulate the amount of marijuana that can be possessed and/or cultivated at home.

There are also differences in the legal definition of medical marijuana at the federal and state levels. At the federal level marijuana is listed as schedule I controlled substance along with heroin, LCD, and ecstasy. Schedule I drugs are characterized as illicit drugs that have a high potential for abuse and have no currently accepted medical treatment use. Consequently, the legal limbo subjects state medical marijuana dispensaries to legal sanctions by federal authorities.[1].

There is a legitimate concern that some of the medical marijuana may end up in the illegal marijuana market. Some evidence suggests MMLs have led to lower prices and an increase in the supply of high-quality marijuana (Anderson et al. 2013). The lower prices for illegal marijuana and an increase in its availability may increase marijuana use among adolescents. Legalization could also change youth perception of the negative health effects associated with regular marijuana consumption by sending a mixed message about the social acceptability of marijuana

[1]Although the federal government deprioritized prosecution of dispensaries abiding state laws, this is not always followed and medical marijuana dispensaries are still subject to being raided or closed by the federal government (Anderson et al. 2014; Hudak 2016)

Table 3.1: States with Medical Marijuana Laws

State	Effective Date
Alaska	March 4, 1999
Arizona	April 14, 2011
California	November 6, 1996
Colorado	June 1, 2001
Connecticut	October 1, 2012
Delaware	July 1, 2011
District of Columbia	July 27, 2010
Hawaii	December 28, 2000
Illinois	January 1, 2014
Maine	December 22, 1999
Maryland	June 1, 2014
Massachusetts	January 1, 2013
Michigan	December 4, 2008
Minnesota	May 30, 2014
Montana	November 2, 2004
Nevada	October 1, 2001
New Hampshire	July 23, 2013
New Jersey	October 1, 2010
New Mexico	July 1, 2007
New York	July 5, 2014
Oregon	December 3, 1998
Pennsylvania	May 17, 2016
Rhode Island	January 3, 2006
Vermont	July 1, 2004
Washington	November 3, 1998

Source: http://medicalmarijuana.procon.org/view.resource.php?resourceID=000881#details and Anderson et al. (2014).

consumption. Theoretically, these demand-side factors could encourage youth marijuana consumption.

While there is some debate on the harmful effects of teen marijuana use (Bechtold et al. 2015), existing evidence suggests that persistent youth marijuana use is associated with greater likelihood of adverse physical and mental health effects, including memory impairments, cognitive impairment, and altered brain development (Meier et al. 2012; Volkow et al. 2014; Hasin et al. 2015). In addition, marijuana

can be addictive, especially for those who begin using it in their teens (Volkow et al. 2014; NIDA 2016).

3.2.2 Studies on MMLs and Teen Marijuana Use

Whether MMLs could lead to higher illegal youth marijuana consumption is hotly debated in the literature. Recently, there has been a renewed interest in examining the relationship between medical marijuana policy and teenage recreational marijuana use as more states consider legalizing medical and/or recreational marijuana. Using YRBS, National Longitudinal Survey of Youth 1997, and Treatment Episode Data, Anderson et al. (2014) finds no evidence legalization of medical marijuana encouraged increased marijuana use among high school students. Similarly, Hasin et al. (2015), using a comprehensive data from Monitoring the Future survey, finds no evidence to support the hypothesis adolescent use of marijuana increased in the year or the following two years after passage.

Several studies explore the relationship between MMLs and marijuana use both at the extensive and intensive margins. Wen et al. (2014), drawing on data from NSDUH, finds evidence MMLs are associated with increased probability and frequency of marijuana use for those 21 and older, but not for those 12-20 years old. They also note that passage of MMLs led to 5-6 percentage point increase in the probability of first time marijuana use among the latter group.

Pacula et al. (2015) emphasizes the importance of considering policy dimensions of MMLs. Specifically, they find that when MMLs provide legal protection for marijuana dispensaries to operate within the state, it leads to an increased illegal consumption of marijuana both among adults and those under 21 years old. In contrast, MMLs with simple medical allowances and patient registration requirements are found to have a negative relationship with recreational marijuana use.

Some studies focus on joint consumption of marijuana with other substances (Dinardo and Lemieux 2001), while others study the effect of marijuana decriminalization

on marijuana consumption (Williams and Bretteville-Jensen 2014, Damrongplasit et al. 2010).[2] Specifically, Dinardo and Lemieux (2001) find that increasing the minimum drinking age, although it reduced the prevalence of alcohol, it also unintentionally led to higher prevalence of marijuana use among high school seniors. Williams and Bretteville-Jensen (2014) find that in the short-run those who would have begun using marijuana in adulthood without decriminalization policy are likely to initiate at younger ages under decriminalization. Damrongplasit et al. (2010) find a significant increase in the probability of marijuana smoking for people living in a decriminalized state.

As more states consider liberalizing their marijuana policies, studying the policy effects on marijuana use is important. In this paper we look at the effect of MMLs on youth marijuana use frequency. This age group is more susceptible to potential harmful health effects from marijuana use. Further, since negative health effects are associated with regular marijuana use by teens, it is important to study the heterogeneity of responsiveness to MMLs by regular vs. light marijuana smokers.

3.3 Data

We use the individual-level youth risk behavior survey (YRBS) data sets for the years 1991 through 2013. The YRBS are school-based surveys where students are asked to fill out self-administered questionnaires. The surveys are conducted every 2 years and provide data on health risk behaviors of 9th through 12th grade students. The national YRBS surveys were designed to obtain a nationally representative sample of 9th through 12th grade students in public and private schools, while state YRBS were designed to be representative of each participating state. For this reason and because state YRBS data is more comprehensive, we draw our data from state YRBS.

[2]Decriminalization of marijuana possession means the person will not be arrested or required to serve prison time for the first time possession of small amounts of marijuana.

Since 1991 a total of 46 states participated in state YRBS, although not every state has participated in every survey year. In 1991 only 9 states participated in state YRBS surveys; this number grew to 32 states by 2003 and 42 states by 2013. Of the 46 states that participated in state YRBS in some years, 37 states made their data publicly available, while 9 states (Colorado, Delaware, Florida, Indiana, Iowa, Massachusetts, Ohio, Pennsylvania, and South Dakota) restrict public sharing of their data sets. We were able to collect restricted use state YRBS data from 8 of these states, bringing the total number of states represented in our combined data set to 45.

There are limitations of YRBS data. As with any school-based survey, the YRBS survey data does not apply to youth who dropped out and do not attend schools. In 2012 the national dropout rates, including youths who may have never attended school, was 2.2% for 16-year olds and 3.5% for 17-year olds.[3] Evidence suggests youth who do not attend schools are more likely to engage in risky behaviors.[4] While alternative surveys, in particular household-based surveys, have the advantage of collecting information on those who dropped out of school, adolescents may be reluctant to provide honest answers, especially when the parent is present at home.

We use frequency of marijuana smoking in the past 30 days as a dependent variable. The frequency of marijuana use in the past 30 days is given in ranges: 0 times, 1-2 times, 3-9 times, 10-19 times, 20-39 times, and 40+ times. Following the related literature, we combine the frequency of marijuana smoking in the past 30 days with demographic variables (age, gender, grade, race) and state-level control variables (median household income, state unemployment rate, state beer excise tax rates, binary indicator for state medical marijuana laws (MML), binary indicators for whether or not states decriminalized marijuana possession, and a binary indicator for

[3]Stark, P., and Noel, A.M. (2015). Trends in High School Dropout and Completion Rates in the United States: 1972-2012 (NCES 2015-015). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved November 27, 2015 from http://nces.ed.gov/pubsearch.

[4]Centers for Disease Control and Prevention. Methodology of the Youth Risk Behavior Surveillance System-2013. MMWR 2013;62(No. RR-1):1-20.

the presence of 0.08 blood alcohol concentration policy). The state-level median household income and beer taxes are converted to 2013 U.S. dollars using the consumer price index for urban consumers.

These data sets are taken from a variety of sources. The median household income is obtained from the U.S. Census Bureau, consumer price index and state unemployment rate are taken from the U.S. Bureau of Labor Statistics, information on state medical marijuana laws are from Procon.org and Anderson et al. (2014). Data on states where marijuana use is decriminalized is taken from Pacula et al. (2003), state beer taxes and 0.08 BAC (blood alcohol concentration) laws come from the Alcohol Policy Information system.[5]

Table 3.1 shows states that enacted MMLs and their effective dates. As of 2013 (the last year in our data), there were 19 states with medical marijuana laws. Since the national and almost all of the state YRBS surveys are conducted during February-May of each odd-numbered year, we set the binary indicators for MML, marijuana decriminalization, BAC laws to 1 if the law became effective during January-April of that year. If the law was effective beginning in May-December of the survey year, we set the binary indicators to 0 for that year and 1 for the next period.

We delete observations if a state with MML lacks data before and after the MML was introduced.[6] In the final combined state YRBS data set, we have slightly more than 900,000 observations. To our knowledge, this is the most comprehensive YRBS data set to date used in studying the effect of MMLs on adolescent marijuana use.

Table 3.2 presents the summary statistics for the variables included in our estimates. Since the responses to marijuana use in the past 30 days are given in intervals, we coded the intervals from 1 to 6, with 1 corresponding to no marijuana use and 6 corresponding to maximum reported range of marijuana use (40+ times).

[5]Retrieved on November 30, 2015 from http://alcoholpolicy.niaaa.nih.gov/.

[6]For example, in state YRBS data sets Colorado has no data before 2005 when the Colorado MML became effective. Consequently, we removed observations for Colorado from the final data set.

Table 3.2: Summary statistics

Statistic	Mean	St. Dev.	Min	Max
MarijNum30	1.51	1.22	1	6
mml	0.17	0.38	0	1
md	0.21	0.40	0	1
rbeer_tax	0.35	0.35	0.00	1.80
Age	15.96	1.23	12	18
Male	0.48	0.50	0	1
linc 2	10.87	0.14	10.49	11.24
unemp	6.06	1.99	2.60	13.70
White	0.62	0.49	0	1
Black	0.14	0.35	0	1
Hispanic	0.13	0.33	0	1
OtherRace	0.11	0.32	0	1
Grade	10.35	1.10	9	12

Number of observations=907,972

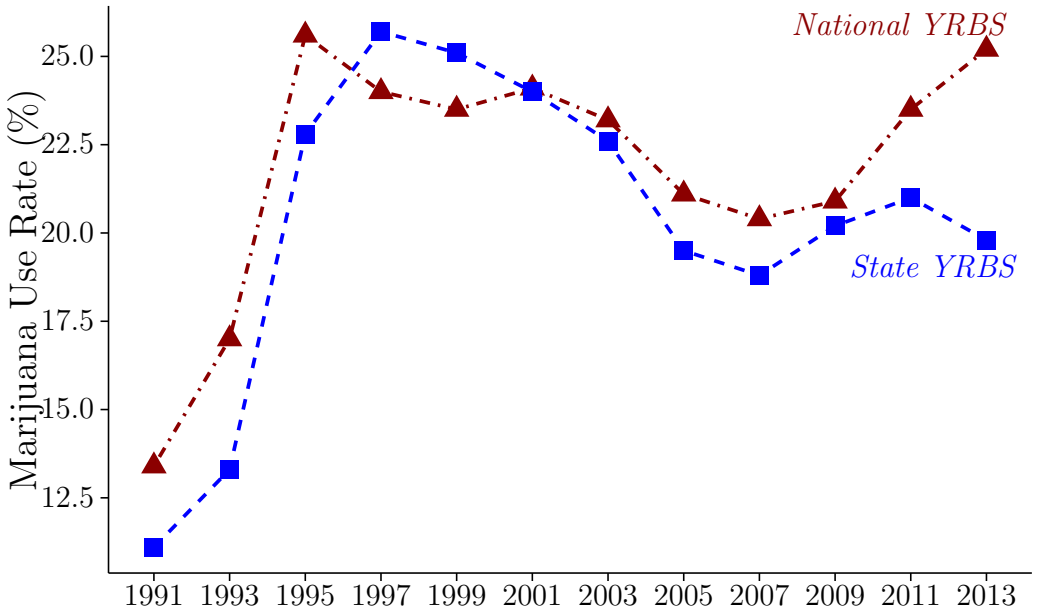

Figure 3.1: Trends in Current Marijuana Use by High School Students, 1991-2013

Notes: Percentage of high school students who used marijuana in the past 30 days. National YRBS is a survey conducted by CDC and State YRBS is a survey conducted by state agencies. National YRBS is based on the national survey conducted by CDC and is representative of 9th through 12th grade students in the U.S. State YRBS is based on unweighted data from the state surveys conducted by state and local agencies. It is representative of students in each state.

Figure 3.1 plots the percentage of 9th through 12th grade students reporting current use of marijuana (any use in the past 30 days). After a downward trend in marijuana consumption rates through mid-2000s, we observe an increase in the marijuana use rate in both the national and state YRBS data. One may think this increase in marijuana use rates could be due to states with MMLs. In order to see a better picture of marijuana use patterns, we plot marijuana consumption rates separately for states with MMLs and states without such laws in place. The survey results indicate that prevalence of youth marijuana consumption is higher in states that passed MMLs (Figure 3.2). This differential pattern holds for every year in our estimation sample. If we consider that the earliest MML was passed in California in 1996, followed by Oregon and Washington in 1998, states with MMLs already had higher marijuana use prevalence among high school students in the years preceding

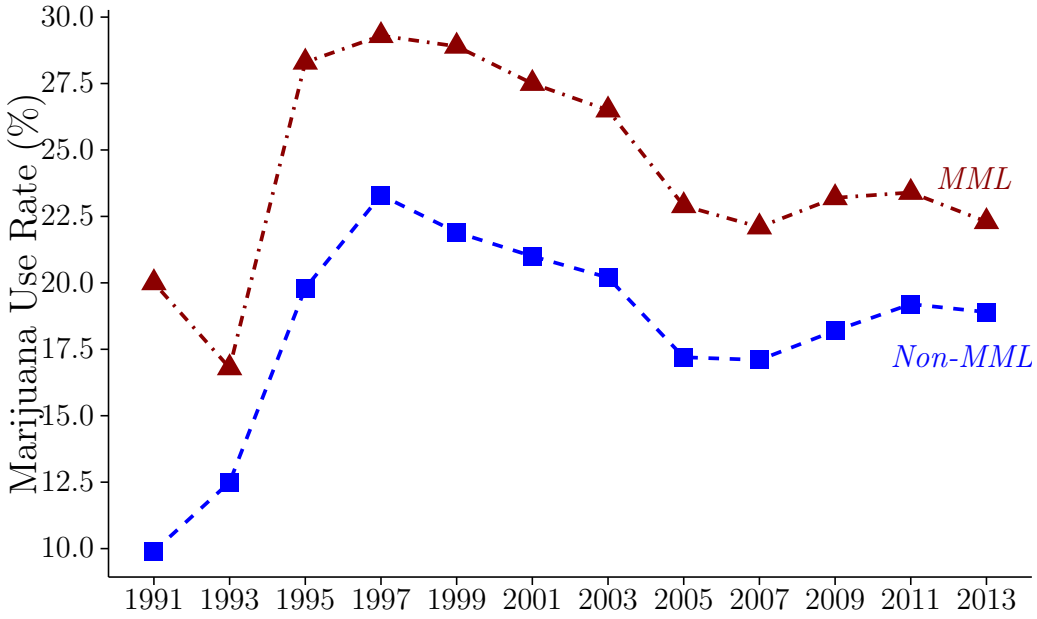

Figure 3.2: Marijuana Use by High School Students in MML States vs. Non-MML States Based on State YRBS, 1991-2013

Notes: Percentage of high school students who used marijuana in the past 30 days in states that enacted MMLs vs. in states that that have not yet enacted MMLs. The figure is based on unweighted State YRBS data.

medical marijuana legalization. Therefore, it is difficult to draw conclusions about the relationship between MMLs and marijuana use by teenagers from only observing trends. Other factors could have influenced the marijuana smoking decisions by adolescents, such as lower perception of risk in states with MMLs.

There are also important differences in the prevalence of teen marijuana consumption by grade level (Figure 3.3). In 2013 about 14% of 9th graders reported smoking marijuana in the past 30 days compared to about 26% for 12th grade students. The lower prevalence rate of illegal marijuana use by freshmen is explained by 9th grade students having lower exposure to illicit drugs.

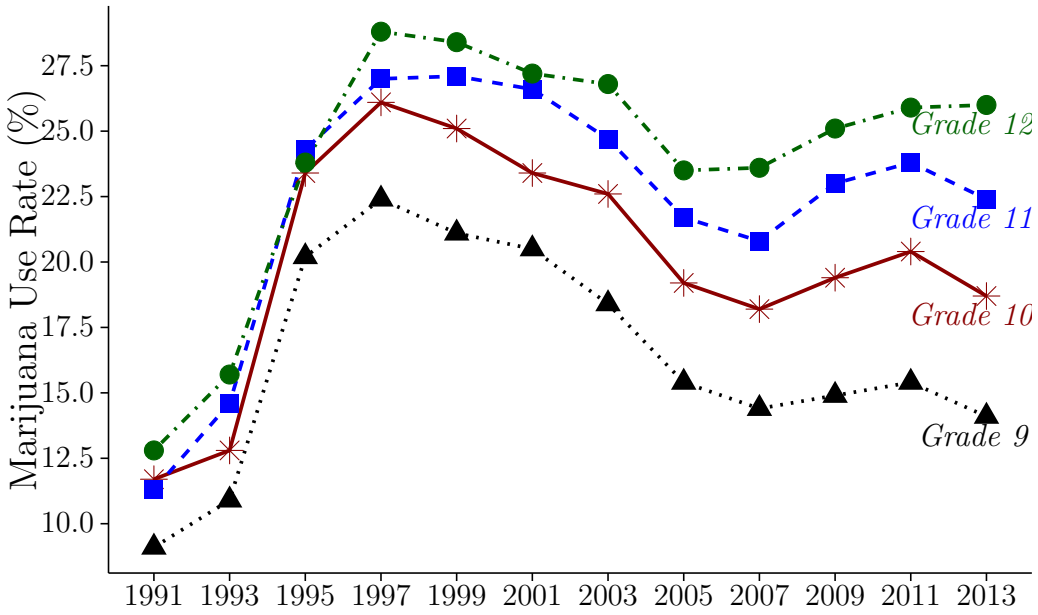

Figure 3.3: Marijuana Use by High School Students by Grade Level Based on State YRBS, 1991-2013

Notes: Percentage of 9th through 12th grader students who used marijuana in the past 30 days. The figure is based on unweighted State YRBS data.

3.4 Econometric Specification

We follow the standard econometric model of marijuana smoking prevalent in the empirical literature. The main econometric specification is as follows:

$$M_{ist} = \beta_0 + \beta_1 X_{ist} + \beta_2 Z_{st} + \beta_3 MML_{st} + \gamma_1 State_i + \gamma_2 Year_t + \epsilon_{ist} \tag{3.1}$$

M_{ist} is the frequency of marijuana smoking in the past 30 days for an individual residing in state s in year t. X_{ist} is a vector of individual characteristics that includes age, gender, grade, and race. Z_{st} is a vector of state-level control variables for state s in year t. Following related literature, we include median household income, unemployment rate, beer excise tax rates, the presence of blood alcohol concentration

(BAC) 0.08 laws, and whether a state has decriminalized the use of marijuana. Median household income and beer excise tax rates have been deflated to 2013 U.S. dollars, the last period in our dataset. The error term is given by ϵ_{ist}.

Our key variable of interest is MML_{st}, a binary indicator for the presence of a medical marijuana law in state s in year t. $State_i$ and $Year_t$ are state and year fixed effects, respectively. State fixed effects absorb any state-specific characteristics that do not vary over time and that are common across treatment and control periods. The year fixed effects can be interpreted as controlling for national level factors that influence youth marijuana use in all states.

The possible answers to marijuana smoking outcome are given in ranges of 0 times, 1-2 times, 3-9 times, 10-19 times, 20-39 times, and 40+ times. We convert these ranges into an ordinal response variable from 1 to 6. For example, 1 refers to nonsmokers, 2 corresponds to those who smoked 1-2 times, 3 corresponds to those who smoked 3-9 times, and so on. Therefore, outcome variables 2 and 3 represent light marijuana users, 5 and 6 represent regular marijuana users (those who smoke almost daily or more). We are interested in estimating the influence of MML on the intensive margin of youth marijuana use. By using marijuana consumption intensity, rather than participation in marijuana smoking, we want to identify the responsiveness of light vs. regular marijuana smokers to MML passage. Since youth marijuana smoking is associated with adverse later life effects, especially for regular smokers, exploring policy effects along the intensive margin of marijuana smoking is of particular importance to policymakers.

We use a quantile regression to estimate the policy responsiveness of marijuana consumption along its intensive margin. Unlike an OLS estimator that measures the mean policy effect on the outcome variable, quantile estimators measure the responsiveness of the outcome variable's conditional distribution. Of particular importance are policy effects at lower and upper quantiles of the conditional distribution of the outcome variable. We hypothesize that MMLs could differentially impact regular marijuana users relative to light marijuana users.

However, quantile regression technique requires the dependent variable to have a continuous distribution. In addition, it is not optimal for count or ordered data such as ours. Therefore, we cannot apply the standard quantile regression technique directly. Instead, we use transformed ordinal quantile regression (TORQUE) method developed by Hong and He (2010). TORQUE achieves a smooth distribution by jittering the ordered dependent variable whose quantiles have a one to one relationship with quantiles of the original dependent variable. The method is based on the data-smoothing technique of Machado and Silva (2005) who suggested adding a uniformly distributed noise to the dependent variable. However, unlike Machado and Silva (2005), Hong and He (2010) do not assume a specific functional form for the link function $\Lambda(\tilde{y})$. Further, although one can apply ordered probit or ordered logit estimators when the outcome variable is an ordered variable, these methods have limitations too. For example,the ordered probit method requires the latent variable be normally distributed. Deviation from normality can potentially affect its accuracy. This is especially problematic given the skewed distribution of marijuana smoking data. In addition, the ordered probit produces estimators in terms of predicted probabilities, so they do not have similar interpretation to quantile or OLS estimators.

The use of TORQUE has at least three advantages. First, the estimation does not rely on any parametric specification of the conditional distributions and therefore it is robust to model misspecification. Second, unlike ordered probit or logit models, predicted intervals from the transformed quantile regression are more informative. For example, MMLs may have differential impacts on the intensive margins of marijuana use. The difference between, say, nonsmokers and light smokers (1 or 2 times) may not be the same as the difference between everyday smokers (20-39 times) and heavy smokers (40+ times). This may be of particular importance for public health officials if the responsiveness to MMLs is different across groups of individuals with various smoking intensity. On the other hand, OLS, ordered probit or ordered logit estimators are conditional mean estimators and do not provide a richer picture of the conditional distribution of the dependent variable. Finally, TORQUE is more robust

than quantile for count estimators due to Machado and Silva (2005) because it does not assume any functional form for the link function. Values of $\Lambda(\tilde{y})$ are estimated with a nonparametric rank-based estimator. The method is not without its weakness. If the relationship between $\hat{\Lambda}$ and $\tilde{y}$ is a complex function, then it could be impossible to get a closed form solution to find a scaling coefficient.

3.5 Results

3.5.1 The Analysis of the Link Function Λ

Figure 3.4 plots the estimated $\Lambda(\tilde{y})$. The function roughly approximates a piecewise linear function with three parts. The slope of $\Lambda(\tilde{y})$ is greatest for $\tilde{y} < 2$ and smallest for $2 < \tilde{y} < 6$. A steeper slope implies the decision to smoke marijuana is harder relative to when the line segment has a flatter slope. This suggests that starting marijuana consumption is a tougher call for nonsmokers. Note that $\Lambda(\tilde{y})$ becomes steeper at $\tilde{y} > 6$ providing evidence that the decision to smoke the next marijuana is

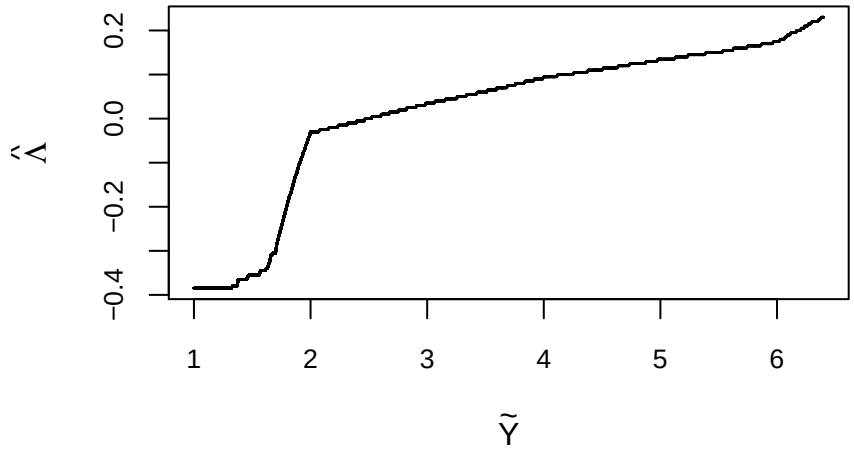

Figure 3.4: Estimation results for Lambda-Marijuana

harder for heavy smokers compared to that of (almost) daily smokers. One possible explanation for this is heavy smokers could switch to stronger drugs. However, we are cautious about this interpretation since we do not have enough evidence to support our claim.

Because we are interested in the marginal effects on the original dependent variable (i.e. the frequency of marijuana smoking), we use quantile regression's invariance to monotone transformation property which allows us to recover the TORQUE estimators. To do so, we fit a piecewise linear function to $\tilde{y}$ on the dependent variable Λ. So we regress $\hat{\Lambda}_i$ on $\tilde{Y}_i$ for $\tilde{Y}_i < 2$ and $\hat{\Lambda}_i = a + c\tilde{Y}_i$ for $\tilde{Y}_i < 1$. Next, by rearranging we get: $\tilde{Y} = \frac{\hat{\Lambda}}{\hat{c}} - \frac{\hat{a}}{\hat{c}}$. Similarly, we repeat the same steps for $2 < \tilde{y} < 6$ and $\tilde{Y}_i > 6$ as well. We refer the interested reader to Hong and He (2010) for details about obtaining the fitted values of $\hat{\beta}_\tau$.

3.5.2 Main Results

Our main results are reported in Table 3.3. We report estimates both for TORQUE and the standard quantile regression of Koenker and Basset (1978). The columns refer to the number of times marijuana is smoked in past month. Our main focus, however, is on the TORQUE estimates. We can think of these estimates as the distance of moving away from own category. For example, a positive sign indicates the distance of moving away to the next, more intense category, while a negative sign means the distance of moving to a previous category (i.e consuming less marijuana than before). We find MMLs have a negative and statistically significant effect on the frequency of marijuana smoking for all groups, especially at lower tails, i.e. for those who are light smokers.

The estimated negative effect of MMLs on teen marijuana consumption may seem counterintuitive at first if one believes that MMLs provide easier access to marijuana. However, we attribute the strong negative effect, particularly at the lower tails of the distribution, to a psychological effect of MMLs. MMLs possibly reduce students'

Table 3.3: TORQUE model results of predicted marijuana use

	Dependent variable: Frequency of marijuana smoking									
	TORQUE					*Standard QR*				
	(0)	(1-2)	(3-9)	(10-19)	(20+)	(0)	(1-2)	(3-9)	(10-19)	(20+)
MML	$-.16^{***}$	$-.08^{***}$	$-.07^{***}$	$-.09^{**}$	$-.01^{*}$	$-.13^{**}$	$-.12^{***}$	$-.03$	$-.04$	$-.03$
	(.039)	(.028)	(.027)	(.043)	(.007)	(.055)	(.026)	(.028)	(.034)	(.02)
MD	$-.06^{***}$	$-.5^{***}$	$-.55^{***}$	$-.42^{**}$	$-.07^{*}$	.004	$-.59^{***}$	$-.78^{***}$	$-.8^{***}$	$-.55^{**}$
	(.021)	(.191)	(.208)	(.192)	(.042)	(.023)	(.128)	(.168)	(.148)	(.242)
beer tax	$-.12^{***}$	$-.87^{***}$	$-.91^{***}$	$-.8^{***}$	$-.14^{***}$	$-.06^{*}$	$-.65^{***}$	$-.98^{***}$	-1.11^{***}	$-.85^{***}$
	(.021)	(.124)	(.135)	(.138)	(.031)	(.034)	(.075)	(.118)	(.095)	(.195)
age	$.08^{***}$	$.45^{***}$	$.41^{***}$	$.39^{***}$	$.06^{***}$	$.06^{**}$	$.27^{***}$	$.37^{***}$	$.34^{***}$	$.25^{***}$
	(.002)	(.013)	(.012)	(.011)	(.003)	(.025)	(.009)	(.013)	(.015)	(.01)
male	$.13^{***}$	$.98^{***}$	1.13^{***}	1.2^{***}	$.23^{***}$	$.38^{***}$	$.76^{***}$	1.16^{***}	1.27^{***}	1.1^{***}
	(.002)	(.014)	(.013)	(.019)	(.003)	(.048)	(.013)	(.014)	(.018)	(.025)
income	.01	.07	.22	.11	.00	$-.08$	$-.02$	$-.27^{*}$	.15	$.29^{*}$
	(.033)	(.168)	(.167)	(.159)	(.038)	(.057)	(.128)	(.155)	(.127)	(.144)
unempl. rate	$.01^{***}$	$.022^{***}$	$.018^{**}$	.01	.00	.00	$.01^{**}$	$-.002$	.01	$.02^{**}$
	(.001)	(.008)	(.009)	(.013)	(.002)	(.003)	(.007)	(.009)	(.01)	(.008)
Race dummies (reference group: black)										
Hispanic	$.03^{***}$	$.2^{***}$	$.24^{***}$	$.28^{***}$	$.07^{***}$	$.05^{*}$	$.14^{***}$	$.25^{***}$	$.22^{***}$	$.21^{***}$
	(.006)	(.027)	(.025)	(.045)	(.007)	(.028)	(.026)	(.028)	(.025)	(.027)
White	$-.05^{***}$	$-.27^{***}$	$-.21^{***}$	$-.2^{***}$	$-.17^{***}$	$-.05^{**}$	$-.21^{***}$	$-.18^{***}$	$-.18^{***}$	-0.12^{***}
	(.004)	(.026)	(.019)	(.03)	(.032)	(.02)	(.017)	(.016)	(.023)	(.018)
Other race	$.03^{***}$	$.25^{***}$	$.4^{***}$	$.46^{***}$	$.1^{***}$	.03	$.18^{***}$	$.37^{***}$	$.37^{***}$	$.3^{***}$
	(.005)	(.031)	(.028)	(.042)	(.007)	(.023)	(.027)	(.026)	(.026)	(.024)
Grade dummies (reference group: 9th grade)										
10th grade	$.02^{***}$	$.14^{***}$	$.11^{***}$	$.12^{***}$	$.02^{***}$	-0.01	.01	$.04^{*}$	$.11^{***}$	$.19^{***}$
	(.002)	(.021)	(.026)	(.027)	(.005)	(.004)	(.009)	(.022)	(.024)	(.026)
11th grade	$.02^{***}$	$.12^{***}$	$.11^{***}$	$.15^{***}$	$.04^{***}$	$.27^{***}$	$.06^{**}$	$.08^{***}$	$.18^{***}$	$.26^{***}$
	(.004)	(.033)	(.032)	(.027)	(.008)	(.096)	(.027)	(.03)	(.037)	(.022)
12th grade	.002	.02	$.09^{*}$	$.19^{***}$	$.06^{***}$	$.41^{***}$	$.12^{***}$	$.14^{***}$	$.18^{***}$	$.25^{***}$
	(.006)	(.046)	(.045)	(.043)	(.011)	(.043)	(.03)	(.049)	(.044)	(.033)
Pseudo-R squared	.026	.03	.033	.033	.037	.043	.061	.076	.082	.069
Observations	907,972	907,972	907,972	907,972	907,972	907,972	907,972	907,972	907,972	907,972

*p<0.1; **p<0.05; ***p<0.01

Notes: SEs for TORQUE estimators are obtained with Delta method, quantile regression standard errors are bootstrapped. Dummy for stated where marijuana is decriminalized is represented by MD.

excitement to try smoking marijuana. Alternatively, it could be medical marijuana increases the availability of high-grade marijuana and, therefore, a lesser amount of it is needed to get the same level of pleasure as that from a larger amount of regular marijuana. There seems to be some support for the latter claim. (Anderson et al. 2013) finds evidence that MMLs have led to lower prices and an increase in the supply of high-quality marijuana.

Similar to MMLs, decriminalization of marijuana has a significantly negative effect at all levels of marijuana smoking frequency, ranging from -0.06 to -0.55. However, unlike the MML effect, decriminalization has a larger impact on moderate marijuana smokers. Similarly, we find beer excise taxes lower teen marijuana consumption.

Upper grade students are estimated to smoke more after MML implementation. Marijuana use also has a positive relationship with gender. Specifically, males who are unemployed are more likely to be positively affected by the presence of MMLs, especially those at the lower tails of the marijuana use distribution. Income, on the other hand, does not have a significant effect on the intensity of smoking. Relative to blacks students, Hispanic and other non-whites students have a higher probability of smoking across all categories.

These results imply strong heterogeneity of MML effects across different categories of marijuana users. Since the most negative effects of marijuana smoking is associated with heavy smokers, the results show they may be very immune to various factors. The most sensitive categories are light to moderate marijuana smokers.

The last five columns report standard quantile regression estimates. They also show negative relationship between MMLs and teens' smoking frequency. However, they are statistically insignificant from moderate to heavy smokers. Although the results from standard quantile regression have consistently same signs as the results from TORQUE, their magnitudes differ. This is clearly seen in the case of nonsmokers and heavy smokers.

We also tested the consistency of our results with those from OLS and ordered probit estimators. Results are presented in Table 3.4. OLS estimates explain average

Table 3.4: OLS and ORDERED PROBITS model results of predicted marihuana use

	Dependent variable: Frequency of marijuana smoking	
	OLS	*Ordered Probits*
MML	$-.34^{***}$	$-.03^{***}$
	(.006)	(.006)
MD	$-.07^{**}$	$-.8^{**}$
	(.034)	(.041)
beer tax	$-.15^{***}$	$-.18^{***}$
	(.025)	(.03)
age	$.09^{***}$	$.1^{***}$
	(.002)	(.002)
male	$.22^{***}$	$.22^{***}$
	(.003)	(.003)
income	$-.02$	$.01^{***}$
	(.031)	(.037)
unemp	.002	$.01^{***}$
	(.002)	(.002)
Race dummies (reference group: black)		
Hispanic	$.04^{***}$	$.5^{***}$
	(.005)	(.006)
White	$-.06^{***}$	$-.07^{***}$
	(.004)	(.005)
Other race	$.05^{***}$	$.05^{***}$
	(.005)	(.006)
Grade dummies (reference group: 9th grade)		
10th grade	$.02^{***}$	$.05^{***}$
	(.004)	(.005)
11th grade	$.02^{***}$	$.05^{***}$
	(.005)	(.006)
12th grade	$.02^{***}$	$.03^{***}$
	(.005)	(.008)
R squared	.03	.019
Observations	907,972	907,972

*p<0.1; **p<0.05; ***p<0.01

marginal effects on the frequency of marijuana smoking. Statistical significance and the direction of the effects are consistent with the resuts from TORQUE model. MMLs have statistically significant negative effect on youth marijuana use frequency. however, we cannot directly compare the coefficient estimates with the estimates from TORQUE since OLS produces only the mean effect. Similarly, the ordered probit produces average effects. Again, we find MMLs have a negative effect on marijuana use by teens, but the effect is much smaller than those found from TORQUE.

3.6 Conclusions

In this paper we study the effect of state medical marijuana laws on youth marijuana consumption frequency. Recently, there has been a renewed interest in studying the impact of MMLs on youth marijuana consumption. Some argue that MMLs increase the supply of marijuana and send a mixed message that use of marijuana is socially acceptable. It may also change youth perception about the health effects of consuming marijuana. Therefore, there is a need for a more rigorous estimation of this relationship given that existing studies have produced mixed results.

In this paper we conducted a more rigorous analysis of the relationship of MMLs with youth marijuana use. In particular, we are interested in disentangling the effects of MMLs on youth marijuana use along its intensive margin. To do so we apply TORQUE, a technique developed by Hong and He (2010), to comprehensive Youth Risk Behavior Survey data set for 1991-2013. TORQUE allows us to estimate the impact of medical marijuana policy on the frequency of marijuana smoking. We find a negative and statistically significant relationship between MMLs and youth marijuana use. This finding is consistent across all specifications. In particular, we find that the effect is stronger for nonsmokers and light smokers. We attribute the negative effects of MMLs on the intensity of youth marijuana smoking to reduced excitement about marijuana among youth in MML states. Perhaps, the saying "forbidden fruit is sweeter" applies to marijuana use by high school students as well. Alternatively,

this effect could be due to the availability of higher-quality marijuana in MML states and, consequently, lesser amounts of it being needed to get the same level of pleasure from consuming regular marijuana.

Importantly, we find the effect of MMLs differs across conditional distribution of marijuana smoking. MMLs have the largest negative effect on those who currently do not use marijuana. This may of particular importance from a public health perspective: current marijuana nonsmokers are less likely to become marijuana smokers with the passage of MMLs. In contrast, marijuana consumption decisions by those who smoke it regularly are less likely to be affected by whether or not a state enacts MML.

Bibliography

Aaronson, D., Lange, F., and Mazumder, B. (2014). Fertility transitions along the extensive and intensive margins. *American Economic Review*, 104(11):3701–24.

Aaronson, D. and Mazumder, B. (2011). The Impact of Rosenwald Schools on Black Achievement. *Journal of Political Economy*, 119(5):821 – 888.

Anderson, D. M., Hansen, B., and Rees, D. I. (2013). Medical marijuana laws, traffic fatalities, and alcohol consumption. *Journal of Law and Economics*, 56(2):333–369.

Anderson, D. M., Hansen, B., and Rees, D. I. (2014). Medical marijuana laws and teen marijuana use. *NBER Working Paper No. 20332*.

Avitabile, C., Clots-Figueras, I., and Masella, P. (2014). Citizenship, fertility, and parental investments. *American Economic Journal: Applied Economics*, 6(4):35–65.

Bailey, M. J. and Collins, W. J. (2011). Did improvements in household technology cause the baby boom? evidence from electrification, appliance diffusion, and the amish. *American Economic Journal: Macroeconomics*, 3(2):189–217.

Baudin, T., de la Croix, D., and Gobbi, P. (2012). DINKs, DEWKs & Co. Marriage, Fertility and Childlessness in the United States. Discussion Papers (IRES - Institut de Recherches Economiques et Sociales) 2012013, Universit catholique de Louvain, Institut de Recherches Economiques et Sociales (IRES).

Baudin, T., de la Croix, D., and Gobbi, P. (2015a). Development policies when accounting for the extensive margin of fertility. Discussion Papers (IRES - Institut de Recherches Economiques et Sociales) 2015003, Universit catholique de Louvain, Institut de Recherches Economiques et Sociales (IRES).

Baudin, T., de la Croix, D., and Gobbi, P. E. (2015b). Fertility and childlessness in the united states. *American Economic Review*, 105(6):1852–82.

Baughman, R. and Dickert-Conlin, S. (2003). Did expanding the eitc promote motherhood? *American Economic Review*, pages 247–251.

Bechtold, J., Simpson, T., White, H. R., and Pardini, D. (2015). Chronic adolescent marijuana use as a risk factor for physical and mental health problems in young adult men. *Psychology of Addictive Behaviors*, 29(3):552.

Becker, G. and Lewis, H. G. (1973). On the interaction between the quantity and quality of children. *Journal of Political Economy*, 81(2):S279–88.

Becker, G. S., Murphy, K. M., and Tamura, R. (1990). Human Capital, Fertility, and Economic Growth. *Journal of Political Economy*, 98(5):S12–37.

Belloni, A. and Chernozhukov, V. (2011). L1-penalized quantile regression in high-dimensional sparse models. *Ann. Statist.*, 39(1):82–130.

Benjamin, D. J., Heffetz, O., Kimball, M. S., and Rees-Jones, A. (2012). What do you think would make you happier? what do you think you would choose? *American Economic Review*, 102(5):2083–2110.

Binder, M. and Coad, A. (2011). From Average Joe's happiness to Miserable Jane and Cheerful John: using quantile regressions to analyze the full subjective well-being distribution. *Journal of Economic Behavior & Organization*, 79(3):275–290.

Blanchflower, D. G. and Oswald, A. J. (2004). Well-being over time in britain and the {USA}. *Journal of Public Economics*, 88(7?8):1359 – 1386.

Boyer, G. R. (1989). Malthus was right after all: poor relief and birth rates in southeastern england. *The journal of political economy*, 93–114.

Bruni, L. and Stanca, L. (2008). Watching alone: Relational goods, television and happiness. *Journal of Economic Behavior and Organization*, 65(3-4):506–528.

Chen, S. (2002). Rank estimation of transformation models. *Econometrica*, 70(4):pp. 1683–1697.

Cigno, A. and Ermisch, J. (1989). A microeconomic analysis of the timing of births. *European Economic Review*, 33(4):737–760.

Clark, A. E., Frijters, P., and Shields, M. A. (2008). Relative income, happiness, and utility: An explanation for the easterlin paradox and other puzzles. *Journal of Economic Literature*, 46(1):95–144.

Clark, A. E. and Oswald, A. J. (1996). Satisfaction and comparison income. *Journal of Public Economics*, 61(3):359 – 381.

Crump, R., Goda, G. S., and Mumford, K. J. (2011). Fertility and the personal exemption: Comment. *The American Economic Review*, pages 1616–1628.

Damrongplasit, K., Hsiao, C., and Zhao, X. (2010). Decriminalization and marijuana smoking prevalence: Evidence from australia. *Journal of Business & Economic Statistics*, 28(3):344–356.

DiNardo, J. and Lemieux, T. (2001). Alcohol, marijuana, and american youth: the unintended consequences of government regulation. *Journal of health economics*, 20(6):991–1010.

Doepke, M., Hazan, M., and Maoz, Y. D. (2007). The Baby Boom and World War II: A Macroeconomic Analysis. IZA Discussion Papers 3253, Institute for the Study of Labor (IZA).

Dolan, P., Peasgood, T., and White, M. (2008). Do we really know what makes us happy? : a review of the economic literature on the factors associated with subjective well-being. *Journal of economic psychology : research in economic psychology and behavioral economics*, 29(1, (2)):94–122.

Easterlin, R. A. (1995). Will raising the incomes of all increase the happiness of all? *Journal of Economic Behavior and Organization*, 27(1):35–47.

Easterlin, R. A. (2001). Income and happiness: Towards a unified theory. *The Economic Journal*, 111(473):pp. 465–484.

Ferrer-i Carbonell, A. and Frijters, P. (2004). How important is methodology for the estimates of the determinants of happiness?*. *The Economic Journal*, 114(497):641–659.

Ferrer-i Carbonell, A. and Ramos, X. (2014). Inequality and happiness. *Journal of Economic Surveys*, 28(5):1016–1027.

Galor, O. (2012). The demographic transition: Causes and consequences. *Cliometrica*, 6(1).

Galor, O. and Weil, D. N. (2000). Population, technology, and growth: From malthusian stagnation to the demographic transition and beyond. *American Economic Review*, 90(4):806–828.

Gobbi, P. E. (2013). A model of voluntary childlessness. *Journal of Population Economics*, 26(3):963–982.

Graham, C., Eggers, A., and Sukhtankar, S. (2004). Does happiness pay?: An exploration based on panel data from russia. *Journal of Economic Behavior and Organization*, 55(3):319–342.

Greenwood, J., Guner, N., and Knowles, J. A. (2003). More on marriage, fertility, and the distribution of income*. *International Economic Review*, 44(3):827–862.

Greenwood, J., Seshadri, A., and Vandenbroucke, G. (2005). The baby boom and baby bust. *The American Economic Review*, 95(1):pp. 183–207.

Guinnane, T. W. (2011). The historical fertility transition: A guide for economists. *Journal of Economic Literature*, 49(3):589–614.

Hakim, C. (2003). A new approach to explaining fertility patterns: Preference theory. *Population and Development Review*, 29(3):pp. 349–374.

Hansen, C. W., Jensen, P. S., and Lønstrup, L. (2014). Accounting for the fertility transition in the us. *Available at SSRN 2380501*.

Hasin, D. S., Wall, M., Keyes, K. M., Cerdá, M., Schulenberg, J., O'Malley, P. M., Galea, S., Pacula, R., and Feng, T. (2015). Medical marijuana laws and adolescent marijuana use in the usa from 1991 to 2014: results from annual, repeated cross-sectional surveys. *The Lancet Psychiatry*, 2(7):601–608.

Hong, H. G. and He, X. (2010). Prediction of functional status for the elderly based on a new ordinal regression model. *Journal of the American Statistical Association*, 105(491):930–941.

Hong, H. G. and Zhou, J. (2013). A multi-index model for quantile regression with ordinal data. *Journal of Applied Statistics*, 40(6):1231–1245.

Horowitz, J. L. (1996). Semiparametric estimation of a regression model with an unknown transformation of the dependent variable. *Econometrica*, 64(1):103–37.

Hotz, V. J. and Miller, R. A. (1988). An empirical analysis of life cycle fertility and female labor supply. *Econometrica: Journal of the Econometric Society*, pages 91–118.

Jones, L. E. and Tertilt, M. (2008). An economic history of fertility in the us 1826-1960.

Kahneman, D. and Krueger, A. B. (2006). Developments in the measurement of subjective well-being. *Journal of Economic Perspectives*, 20(1):3–24.

Koenker, R. and Bassett, Gilbert, J. (1978). Regression quantiles. *Econometrica*, 46(1):pp. 33–50.

Krueger, A. B. and Schkade, D. A. (2008). The reliability of subjective well-being measures. *Journal of Public Economics*, 92(8?9):1833 – 1845. Special Issue: Happiness and Public Economics.

La Ferrara, E., Chong, A., and Duryea, S. (2012). Soap operas and fertility: Evidence from brazil. *American Economic Journal: Applied Economics*, 4(4):1–31.

Machado, J. and Silva, S. (2005). Quantiles for counts. *Journal of the American Statistical Association*, 100(472):1226–1237.

Meier, M. H., Caspi, A., Ambler, A., Harrington, H., Houts, R., Keefe, R. S., McDonald, K., Ward, A., Poulton, R., and Moffitt, T. E. (2012). Persistent cannabis users show neuropsychological decline from childhood to midlife. *Proceedings of the National Academy of Sciences*, 109(40):E2657–E2664.

Momota, A. (2015). Intensive and extensive margins of fertility, capital accumulation, and economic welfare. *KIER Discussion Paper Series*, 917.

National Institute on Drug Abuse (2016). Marijuana. Retrieved on April 26, 2016.

Oswald, A. J. and Powdthavee, N. (2008). Does happiness adapt? a longitudinal study of disability with implications for economists and judges. *Journal of Public Economics*, 92(5?6):1061 – 1077.

Pacula, R. L., Chriqui, J. F., and King, J. (2003). Marijuana decriminalization: What does it mean in the united states? *NBER Working Paper No. 9690*.

Pacula, R. L., Powell, D., Heaton, P., and Sevigny, E. L. (2015). Assessing the effects of medical marijuana laws on marijuana use: the devil is in the details. *Journal of Policy Analysis and Management*, 34(1):7–31.

Shields, M. P. and Tracy, R. L. (1986). Four themes in fertility research. *Southern Economic Journal*, pages 201–216.

Volkow, N. D., Baler, R. D., Compton, W. M., and Weiss, S. R. (2014). Adverse health effects of marijuana use. *New England Journal of Medicine*, 370(23):2219–2227.

Wen, H., Hockenberry, J., and Cummings, J. R. (2014). The effect of medical marijuana laws on marijuana, alcohol, and hard drug use. *NBER Working Paper No. 20085*.

Whittington, L. A., Alm, J., and Peters, H. E. (1990). Fertility and the personal exemption: implicit pronatalist policy in the united states. *The American Economic Review*, pages 545–556.

Williams, J. and Bretteville-Jensen, A. L. (2014). Does liberalizing cannabis laws increase cannabis use? *Journal of health economics*, 36:20–32.

Winkelmann, L. and Winkelmann, R. (1998). Why are the unemployed so unhappy? evidence from panel data. *Economica*, 65(257):1 – 15.

Vita

Okila Elboeva was born in Shakhrisabz, Uzbekistan. She received her Bachelor Degree in Finance in 1999 from Tashkent Finance Institute. After working at National Bank of Uzbekistan, she joined the PhD program at the University of Tennessee, Knoxville in August 2011. She will receive her doctoral degree in May 2016 and will begin to work as an Assistant Professor of Economics at Lincoln Memorial University in August 2016.