Multiple Fractional Response Variables with Continuous Endogenous Explanatory Variables.

Nam, Suhyeon

October 2012

Multiple Fractional Response Variables with Continuous Endogenous Explanatory Variables

Suhyeon Nam [†‡]

Michigan State University

October 2012

Abstract

Multiple fractional response variables have two features. Each response is between zero and one, and the sum of the responses is one. In this paper, I develop an estimation method not only accounting for these two features, but also allowing for endogeneity. It is a two step estimation method employing a control function approach: the first step generates a control function using a linear regression, and the second step maximizes the multinomial log likelihood function with the multinomial logit conditional mean which depends on the control function generated in the first step. Monte Carlo simulations examine the performance of the estimation method when the conditional mean in the second step is misspecified. The simulation results provide evidence that the method's average partial effects (APEs) estimates approximate well true APEs and that the method's approximations is preferable to an alternative linear method. I apply this method to the Michigan Educational Assessment Program data in order to estimate the effects of public school spending on fourth grade math test outcomes, which are categorized into one of four levels. The effects of spending on the top two levels are statistically significant while almost those on the others are not.

keywords: Multiple fractional responses; Endogeneity; Partial effects; Two step estimation; Control function approach; Misspecified conditional mean; Monte Carlo simulation

---

[†] Department of Economics, Michigan State University, East Lansing, MI 48824, namsu@msu.edu.

[‡] I am very grateful for the guidance and support of my advisor, Jeffrey M. Wooldridge. I would like to thank my dissertation committee members, Peter Schmidt, Leslie Papke and Keneth Frank for their useful advice. I also thank Sungsam Chung, Do Won Kwak and the seminar participants at Michigan State University for their helpful comments. All errors are my own responsibility.

1 Introduction

Fractional responses have interesting functional form issues that have been attracting econometricians' attentions. The research began with a single fractional response, a fractional scalar y_i, which has a salient feature - the bounded nature: $0 \leq y_i \leq 1$. Then it has moved to two kinds of systems of fractional responses. One is panel data setting in which a cross sectional unit has relatively smaller time periods. The other is multiple responses in which a cross sectional unit has a set of several choices, which is the interest of this paper.

As to single fractional responses, it is true that the OLS estimator or the IV estimator of a linear model are consistent even though they ignore the bounded nature. They, however, do not guarantee that their fitted values lie within the unit interval nor that their partial effect estimates for regressors' extreme values are good. Theses are the same drawbacks as the linear probability model for binary response has. The log-odds transformation, $\log\frac{y}{1-y}$, is a traditional solution to recognize the bounded nature. But it requires the responses to be strictly inside the unit interval. Papke and Wooldridge (1996) introduce a quasi maximum likelihood estimation (QMLE), a particular QMLE that Gourieroux, Monfort, and Trognon (1984) describe. This nonlinear estimation method directly models the conditional mean of the responses as an appropriate function. It can provide a consistent estimator even when the responses take the boundary values.

Papke and Wooldridge (2008) extend their single fractional response discussion to panel data with allowing for endogeneity. They allow time invariant unobserved effect to be correlated with explanatory variables and develop another QMLE method employing a control function approach to account for endogeneity.

Multiple fractional responses have one additional feature as well as the bounded nature: an adding-up constraint - the sum of an observation's responses is one. Suppose a researcher studies individual's asset allocation behavior with the research question, how much people have their pension funds invested in stocks and bonds. Let her response variables be an individual i's two shares of pension funds, $(y_{i,stock}, y_{i,bond})$ where $y_{i,stock} + y_{i,bond} = 1$. Since there are only two shares, this example can fall into the single fractional response category. But when there are more than two shares, it requires a different estimation method to exploit the whole available information. For example, there are four shares in my application. I estimate the effects of Michigan's public school spending on fourth grade math test outcomes using the year of 2005 Michigan Educational Assessment Program (MEAP) data. The test outcomes are graded the student's level of proficiency from Level 1 (Excellence) to Level 4 (Apprentice). The MEAP data provides districts' shares of the four levels where each share is a fraction and the sum of each district's four shares is one.

The estimation method by Sivakumar and Bhat (2002) handles multiple fractional responses with more than two shares. It is a method of QMLE with the multinomial distribution and the multinomial logit conditional mean specification. It is a multivariate generalization of the method proposed by Papke and Wooldridge (1996). In the economic literature, Mullahy (NBER 2010) studies the same QMLE method with more details. Buis (2008) writes a STATA® module of this QMLE method and dubs it as "fractional multinomial logit (fmlogit)." In this paper, I also refer this QMLE as fractional multinomial logit or fmlogit.

Although these studies develop a new estimation method for multiple fractional responses which can consistently estimate the parameters in the mean as long as the mean specification is correct, they do not address endogeneity. In empirical works, however, endogeneity arises very prevalently. The asset allocation researcher might suspect that some regressors such as age or income are correlated with individual's risk attitude, which is unobservable to her but definitely affects people's decisions. In my application, I suspect that the district school spending is correlated with unobserved district effects such as parental involvement. This endogeneity issue may lead to inconsistency of the fractional multinomial logit estimation.

Here I develop an estimation method for multiple fractional responses with endogenous explanatory variables. In the model, I allow a continuous endogenous explanatory variable to be correlated with an unobserved omitted variable. To deal with it, I propose a two step estimation method employing a control function approach. The first step generates a control function and the second step applies fractional multinomial logit with including the control function as extra regressors in the conditional mean. It provides consistent estimates of the conditional mean parameters provided that the conditional mean specification in the second step is correct.

A distinct feature of this method is that although the multinomial logit specification for the second step is sensible as a conditional mean for multiple fractional responses, it is not underpinned by the underlying assumptions. The functional form of the conditional mean in the second step is determined by the two elements. One is the functional form of the conditional mean depending on the unobserved omitted variable. The other is the distributional assumption of the error, which appears when control function approach is applied. However, no combination of an explicit functional form and a distribution is known to derive the multinomial logit functional form. Thus, I suggest directly specifying the conditional mean of the second step as multinomial logit without assuming the underlying conditions. This approach reflects the way in which Petrin and Train (2010) generate a mixed logit.

In order to investigate how the two step estimation method works when the multinomial logit specification is wrong, I conduct Monte Carlo simulations. The simulations focus on whether or not the estimation method can approximate well the average partial effects of the endogenous explanatory variable, which is the partial effects of the endogenous explanatory variable on the conditional mean averaged across the population. I compare the method's performance with an alternative linear model's.

The simulation results provide evidence that the two step estimation method approximates well the true APEs as long as a strong instrument is used. It is preferable to the alternative linear approach. With a weak instrument, its approximation is not good. However the approximation by the linear approach is worse.

The rest of the paper is organized as follows. Section 2 describes the model and discusses the two step estimation method with more details. Section 3 presents a Monte Carlo simulation design and results where the conditional mean of the estimation method is misspecified. Section 4 includes an application of the estimation method to obtain the average partial effects estimates of district school spending on fourth grade Michigan Educational Assessment Program math test outcomes. And Section 5 concludes the paper.

2 The Model and Estimation with Endogeneity

I assume that random sampling across the cross section is available, each cross sectional unit i has G choices, and the sum across choices is one. The dependent variable for i is

$$y_i = \begin{pmatrix} y_{i1} \\ \vdots \\ y_{ig} \\ \vdots \\ y_{iG} \end{pmatrix} \tag{1}$$

$$\text{where} \quad 0 \leq y_{ig} \leq 1 \quad \text{and} \tag{2}$$

$$\sum_{g}^{G} y_{ig} = 1. \tag{3}$$

(2) and (3) represent the two features of multiple fractional responses: the bounded nature and the adding-up constraint, respectively. For $X_i = (x_{i1}, \cdots, x_{iG})$, the set of explanatory variables in all choices, I assume the conditional mean as

$$E(y_{ig}|\mathbf{X}_i) = G_g(\boldsymbol{\beta}, \mathbf{X}_i), \qquad g = 1, 2, \cdots, G, \tag{4}$$

$$\text{where} \quad 0 < G_g(\cdot) < 1 \text{ and} \tag{5}$$

$$\sum_{g}^{G} G_g = 1. \tag{6}$$

(5) ensures that the fitted value will lie between zero and one. The adding-up constraint (3) leads to (6). $G(\cdot)$ can be any function satisfying both (5) and (6).[1]

To allow for endogeneity, I assume that $\mathbf{X}_i$ includes a continuous endogenous ex-

[1] Fractional multinomial logit by Sivakmar and Bhat (2002) and Mullahy (NBER 2010) specifies it as multinomial logit probabilities.

planatory variable w_{ig} and an unobserved omitted variable r_{ig}, and that w_{ig} and r_{ig} are correlated. To simplify the exposition, I assume that w_{ig} and r_{ig} are invariant across choice : $\forall g$, $w_{ig} = w_i$, $r_{ig} = r_i$ where w_i and r_i are scalars.

To deal with the endogeneity, I employ a control function approach. It includes extra regressors in the estimating equation so that the remaining variation in the endogenous explanatory variable would not be correlated with unobservable. To accommodate this approach to the model, (4) is now written as

$$E(y_{ig}|\mathbf{X}_i) = E(y_{ig}|\mathbf{Z}_i, w_i, r_i) = G_g(\boldsymbol{\beta}, \mathbf{Z}_{i1}, w_i, r_i), \qquad g = 1, 2, \cdots, G, \tag{7}$$

$\mathbf{Z}_i \equiv (\mathbf{z}_{i1}, \cdots, \mathbf{z}_{iG})$ is the set of exogenous variables in all choices where $\mathbf{z}_{ig} = \left(\mathbf{z}_{i1g} \vdots \mathbf{z}_{i2g}\right)$ is the exogenous variables vector for choice g and a constant is included in $\mathbf{z}_{i1g}$, $\forall g$. $\mathbf{Z}_{i1} \equiv (\mathbf{z}_{i11}, \cdots, \mathbf{z}_{i1G})$ is the set of $\mathbf{z}_{i1g}$. Note that the approach requires an exclusion restriction: only some part of the exogenous variables appear in (7), the conditional mean depending on r_i.

The approach requires additional assumptions. Dropping the cross-sectional unit i, for each g,

$$w_g = w = \mathbf{Z}\boldsymbol{\pi} = \mathbf{Z_1}\boldsymbol{\pi}_1 + \mathbf{Z_2}\boldsymbol{\pi}_2 + v, \tag{8}$$

$$r_g = r = \rho v + e, \tag{9}$$

$$D(e|\mathbf{Z}, v) = D(e). \tag{10}$$

(8) is the reduced form of the endogenous variable w where $\boldsymbol{\pi}' = (\boldsymbol{\pi}_1' \boldsymbol{\pi}_2')$ is the parameter vector and $\mathbf{Z}_{i2} \equiv (\mathbf{z}_{i21}, \cdots, \mathbf{z}_{i2G})$ is the set of excluded exogenous variables. (9) is the linear projection of the omitted variable r on v, the reduced form error in (8), which is

the control function. It reveals that if there is any correlation between w and r, it can only come through v. ρ shows how much w is correlated with r, and consequently tells whether w is endogenous or not. (10) shows that e is independent of $(\mathbf{Z}, v)$.

Then the mean function conditional on the reduced form error v_i, is derived as

$$E(y_{ig}|\mathbf{Z}_i, w_i, v_i) = K_g(\boldsymbol{\theta}, \mathbf{Z}_{i1}, w_i, v_i), \tag{11}$$

$$\text{where} \quad 0 < K_g(\cdot) < 1 \text{ and} \tag{12}$$

$$\sum_g^G K_g = 1. \tag{13}$$

If v_i is observed, $\boldsymbol{\theta}$ can be estimated by nonlinear least squares or a QMLE using multinomial distribution with specifying $K_g(\cdot)$ as a proper functional form. When v_i is unobserved, a simple way to estimate the parameters in (11) is to replace v_i with its consistent estimator and apply one of those two methods. Therefore, I suggest the following two step procedure:

PROCEDURE 2.1

Step 1. Obtain the OLS residual $\hat{v}_i$ from the regression of w_i on $\mathbf{z}_i$.

Step 2. Apply fractional multinomial logit of (y_{i1}, y_{i2}, y_{iG}) on $\mathbf{z}_{i1}$, w_i and $\hat{v}_i$ to estimate $\boldsymbol{\theta}$. This is a QMLE using the following log likelihood and conditional mean.

$$\ell_i(\boldsymbol{\theta}) = \sum_g^G y_{ig} \log K_g(\boldsymbol{\theta}, \mathbf{z}_{i1}, w_i, \hat{v}_i) \text{ and} \tag{14}$$

$$K_g(\boldsymbol{\theta}, \hat{\mathbf{x}}_{vi}) = \frac{\exp\left(\hat{\mathbf{x}}_{vi}\boldsymbol{\theta}_g\right)}{\sum_h^G \exp\left(\hat{\mathbf{x}}_{vi}\boldsymbol{\theta}_h\right)} \tag{15}$$

where $\hat{\mathbf{x}}_{vi} = \begin{pmatrix} \mathbf{z}_{i1} & w_i & \hat{v}_i \end{pmatrix}$ is a $1 \times p$ vector, $\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\theta}_1' & \dots & \boldsymbol{\theta}_G' \end{pmatrix}'$ is a $pG \times 1$ parameter vector and $\boldsymbol{\theta}_1 = \mathbf{0}$[2].

The two step estimation method specifies $K_g(\cdot)$ as multinomial logit since it is a natural choice that satisfies (12) and (13). Hence, this method is appropriate for problems where characteristics of choices are unimportant or are not of interest because conditioning variables in the basic multinomial logit change by unit i, not by choice g. With the specification, (8) becomes

$$w_{ig} = w_i = \mathbf{z}_i\boldsymbol{\pi} = \mathbf{z}_{i1}\boldsymbol{\pi}_1 + \mathbf{z}_{i2}\boldsymbol{\pi}_2 + v_i, \tag{16}$$

where $\mathbf{z}_i = \begin{pmatrix}\mathbf{z}_{i1} & \vdots & \mathbf{z}_{i2}\end{pmatrix}$ is a $1 \times M$ vector of exogenous variables which are the same across g, and then the first step of the procedure comes from (16).

Under the assumption (15), the second step consistently estimates $\boldsymbol{\theta}$. For the consistency, it does not need any additional assumptions. It is because the QMLE uses multinomial distribution, which is a member of linear exponential family (LEF). Gourieroux, Monfort, and Trognon (1984) show that a QMLE with a distribution in LEF provides consistent estimates of the parameters in a correctly specified conditional mean even when the rest of distribution is misspcified. The asymptotic variance of $\widehat{\boldsymbol{\theta}}$ needs to consider the additional variation from the first step. Appendix derives its asymptotic variance estimator.

Notice that the method does not make any assumptions regarding $G_g(\cdot)$ and $D(e)$ even though they determine the functional form of $K_g(\cdot)$. If $K_g(\cdot)$ is explicitly derived by assuming them, $\boldsymbol{\theta}$ is consistently estimated by the two step estimation using the derived form for $K_g(\cdot)$ instead of multinomial logit. However, no combination of $G_g(\cdot)$ and $D(e)$ is known to derive an explicit form of $K_g(\cdot)$.

[2] the first choice is chosen as a reference.

Although a natural choice of $G_g(\cdot)$ is also multinomial logit because it holds (5) and (6), it is not a proper choice since it does not derive an explicit form of $K_g(\cdot)$ whatever $D(e)$ is. $G_g(\cdot)$ specified as the following can bring a known form of $K_g(\cdot)$ with assuming that e is normally distributed.

$$G_g(\boldsymbol{\beta}, \mathbf{X}_i) = \Phi(\mathbf{x}_{ig}\boldsymbol{\beta}) \qquad g = 1, \cdots, G-1,$$
$$G_G(\boldsymbol{\beta}, \mathbf{X}_i) = 1 - \sum_{g}^{G-1} \Phi(\mathbf{x}_{ig}\boldsymbol{\beta}) \tag{17}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function. Based on the mixing property of normal distribution, $K_g(\cdot)$ becomes a similar form of (17). But $G_G(\cdot)$, the conditional mean of the last choice, is not necessarily between zero and one, which violates (5). So (17) is inappropriate as well.

Therefore, I suggest directly specifying $K_g(\cdot)$ as multinomial logit without assuming anything about $G_g(\cdot)$ and $D(e)$. This idea reflects the manner in which Petrin and Train (2010) employ a control function approach. They divide the structural error in their consumer utility into two parts to generate a mixed logit. Without the structural error's distributional assumption, one divided part is assumed to be normal and the other is assumed to be type 1 extreme value.

The consistency of the proposed two step estimation method hinges on the multinomial logit specification which are not supported by underlying assumptions. So in the next section, I conduct Monte Carlo simulations to evaluate the performance of the estimation method when the multinomial logit specification is not true. The simulations examine how well the partial effects estimates approximate the true ones. Especially, I am interested in the partial effects of the endogenous explanatory variable w on the conditional mean, $\frac{\partial E(y_g|\mathbf{z}, w, r)}{\partial w}$. It, however, is not identified due to the unobserved r.

Instead, I use average partial effects (APEs), which can be identified by averaging the partial effects over the distribution of r, $E_{r_i}\left(\frac{\partial E(y_g|\mathbf{z}^\circ, w^\circ, r_i)}{\partial w}\right)$.

3 Monte Carlo Simulation Design

3.1 Data Generating Process

The right-hand side variables:

The number of observation N and the number of iteration are set to be 500 and 1000, respectively. For each replication, I generate 500 observations of $\mathbf{z}_i$, w_i, r_i, v_i and e_i as following.

- $\mathbf{z}_i = \begin{pmatrix} \mathbf{z}_{i1} \vdots z_{i2} \end{pmatrix} = \begin{pmatrix} 1 & z_{i1} & \vdots & z_{i2} \end{pmatrix}_{1\times 3}$

  $$\begin{pmatrix} z_{i1} \\ z_{i2} \end{pmatrix} \sim MV\,Normal\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \tau \\ \tau & 1 \end{pmatrix}\right]$$

  There are one included exogenous variable and one excluded exogenous variable where they are drawn from multivariate normal distribution. Some simulations allow them to be correlated: $\tau \neq 0$.

- $D(e)$ has one of the three distribution.

  (a) $e \sim Normal(0,1)$

  (b) $e \sim Logistic(0,1)$

  (c) $e \sim \chi^2_3$

  To study various misspecifications, three distributions of e are used: two symmetric distributions and one asymmetric distribution.

- $v \sim Normal(0, \sigma^2)$

- $w_i = \pi_1 z_{i1} + \pi_2 z_{i2} + v_i$

  The endogenous variable is generated by (16). The constant coefficient parameter is set to be zero. π_2 shows the instrument's predictive power. Several values of these parameters are used in the simulations. When the parameters have different values, the value of σ^2 is also adjusted for the variance of w_i to be invariant[3].

- $r_i = \rho v_i + e_i$ where $\rho \in \{0.1, 0.5, 1\}$

  The omitted variable is generated by (9). ρ indicates the amount of endogeneity.

The conditional mean specifications:

In the simulation, $G_g(\cdot)$, the conditional mean depending on r_i, is multinomial logit.

$$E(y_{ig}|\mathbf{z}_i, w_i, r_i) = G_g(\boldsymbol{\beta}, \mathbf{z}_{i1}, w_i, r_i) = \frac{\exp\left(\mathbf{x}_i \boldsymbol{\beta}_g\right)}{\sum_h^G \exp\left(\mathbf{x}_i \boldsymbol{\beta}_h\right)} \tag{18}$$

where $\mathbf{x}_i = \begin{pmatrix} \mathbf{z}_{i1} & w_i & r_i \end{pmatrix}$ is a 1×4 vector, $\boldsymbol{\beta} = \begin{pmatrix} \boldsymbol{\beta}_1' & \dots & \boldsymbol{\beta}_G' \end{pmatrix}'$ is a $4G \times 1$ parameter vector and $\boldsymbol{\beta}_1 = \mathbf{0}$ since the first choice is chosen as a reference. The parameters for the other choices are set to be 1 in the simulation: $\boldsymbol{\beta}_g = \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix}'$ for $g = 2, \cdots, G$. Note that (15) is a *wrong* specification under any of the three distributions for e and (18).

The multiple fractional dependent variables:

The number of choice G is chosen as 3. The multiple fractional dependent variables for each observations i are generated by the following process.

1) By using (18) and the variables generated above, calculate the response probabilities, G_{i1}, G_{i2}, and G_{i3}.

[3] $Var(w) = 2$

2) Draw 100 multinomial outcomes among 1, 2, and 3 based on the calculated response probabilities.
3) Count the frequency and obtain the proportion for each outcome.

For instances, if 1 is drawn 50 times, 2 is drawn 30 times, and 3 is drawn 20 times for the observation i, then $y_{i1} = 0.5$, $y_{i2} = 0.3$, and $y_{i3} = 0.2$. Through this process, the upper corner 1 is generated only for the reference choice, which is the first choice in this simulation, while the lower corner 0 is generated for all three choices. It is due to the multinomial logit response probabilities.

3.2 Estimation

After generating the data, the simulations examine if the APEs estimates by the two step estimation method approximate well the true APEs and if the approximation is better than an alternative linear model's.

True APEs

The true APEs of choice g, evaluated at $(\mathbf{z}_1^\circ, w^\circ)$, is

$$E_{r_i}\left[\frac{\partial G_g(\mathbf{x}^\circ)}{\partial w}\right] = E_{r_i}\left[G_{ig}^\circ \cdot \left(\beta_{wg} - \frac{\sum_h^3 \beta_{wh} \exp\left(\mathbf{x}^\circ \boldsymbol{\beta}_h\right)}{\sum_h^3 \exp\left(\mathbf{x}^\circ \boldsymbol{\beta}_h\right)}\right)\right], \tag{19}$$

where $\mathbf{x}^\circ = \begin{pmatrix}\mathbf{z}_1^\circ & w^\circ & r_i\end{pmatrix}$ and $G_{ig}^\circ = G_g(\mathbf{x}^\circ)$.

By the law of large numbers, I obtain the true APEs as following:

$$\frac{1}{N}\sum_i^N \left[G_{ig}^\circ \cdot \left(\beta_{wg} - \frac{\sum_h^3 \beta_{wh} \exp\left(\mathbf{x}^\circ \boldsymbol{\beta}_h\right)}{\sum_h^3 \exp\left(\mathbf{x}^\circ \boldsymbol{\beta}_h\right)}\right)\right] \tag{20}$$

Two step estimation's APEs estimates

The two step estimation's APEs estimate for choice g, evaluated at $(\mathbf{z}_1^\circ, w^\circ)$ is calculated as

$$\widehat{\delta}_{wg} \equiv \frac{1}{N}\sum_{i}^{N}\left[\widehat{K}_{ig}^{\circ}\cdot\left(\widehat{\theta}_{wg} - \frac{\sum_h^3 \widehat{\theta}_{wh}\exp\left(\widehat{\mathbf{x}}_v^\circ \widehat{\theta}_h\right)}{\sum_h^3 \exp\left(\widehat{\mathbf{x}}_v^\circ \widehat{\theta}_h\right)}\right)\right] \tag{21}$$

where $\widehat{\boldsymbol{\theta}}$ is obtained from PROCEDURE 2.1, $\widehat{\mathbf{x}}_v^\circ = \begin{pmatrix}\mathbf{z}_1^\circ & w^\circ & \widehat{v}_i\end{pmatrix}$ and $\widehat{K}_{ig}^\circ = K_g(\widehat{\mathbf{x}}_v^\circ)$. $\widehat{\delta}_{wg}$'s asymptotic standard errors are described in Appendix.

To obtain a single APEs estimate, I use two approaches. One is averaging the estimates out across the sample and the other is evaluating them at a certain set of values, $(\overline{\mathbf{z}}_1, w_p)$ where $\overline{\mathbf{z}}_1$ is the mean of $\mathbf{z}_1$ and w_p stands for the 10^{th}, 25^{th}, 50^{th}, 75^{th} and 90^{th} percentiles of w's distribution. I call the former APEs estimate "average APEs estimate" and the latter "percentile APEs estimate".

Linear model's APEs estimates

Researchers, who are inclined to use a linear model rather than a nonlinear model, would drop one of three choices and apply the linear control function (LCF) approach to the remaining choices. $\widehat{\gamma}_{wg}$, the coefficient estimates of this LCF estimation are comparable to the APEs estimates[4]. The simulations drop the first choice, the reference choice.

PROCEDURE 3.1

Step 1. Obtain the OLS residual $\widehat{v}_i$ from the regression of w_i on $\mathbf{z}_i$.

This is the same step as Procedure 2.1.

Step 2. For each $g = 2, 3$, regress y_{ig} on $\mathbf{z}_{i1}$, w_i and $\widehat{v}_i$ to estimate $\boldsymbol{\gamma}_g$, where $\boldsymbol{\gamma}_g$ is a 4×1 parameter vector for choice g.

[4] Since $\widehat{\gamma}_{wg}$ does not depend on the variable, the LCF approach has one estimate for both the average APEs estimate and the percentile APEs estimate.

The first choice coefficients are obtained by the constraint that the sum of y_{ig} across choices is one : $\boldsymbol{\gamma}_1 = \mathbf{e}_1 - \boldsymbol{\gamma}_2 - \boldsymbol{\gamma}_3$, where $\mathbf{e}_1$ is a 1×4 unit vector. The standard errors of $\widehat{\gamma}_g$ also need the adjustment to take the extra variation from the first step into account: See Appendix.

Additional Estimation

Some simulations examine whether allowing for the flexibility of the control function improves the approximation or not when there exists a nonlinearity of w in the model. I include w^2 in the model[5] and let the second steps in PROCEDURE 2.1 and PROCEDURE 3.1 contain $\widehat{v}_i^2$ and $\widehat{v}_i^3$ as well as $\widehat{v}_i$[6].

3.3 Simulation Results

The first columns of Tables show whether the model includes w^2 or not: w indicates the model including only w, and w^2 indicates the models including w^2 as well as w. The second rows of Tables represent the choice, g. Tables report the means of the estimates over the 1000 replications (Mean), the standard deviations (SD), and the means of the 1000 adjusted standard errors (SE). When the mean of the standard errors is huge, the median of the adjusted standard errors (SE*) is reported. Due to the calculational difficulty, I can not report $\widehat{\delta}_{wg}$'s standard errors: obtaining them in the simulation takes too much time to obtain. The results are rounded off to the three decimal places in Tables.

Condition 1

Table 1 through Table 3 report the results of the simulations where the instrument's predictive power is strong ($\pi_2 = 1$) and the endogeneity is also strong ($\rho = 1$):

[5] The corresponding parameter is set to be 0.1: $\beta_{w^2g} = 0.1$ for $g = 2, 3$.

[6] The LCF APEs estimates, $\widehat{\gamma}_{wg} + w\widehat{\gamma}_{w^2g}$, are comparable to the two APEs estimates.

- $w_i = z_{i2} + v_i$
- $r_i = v_i + e_i$

For those simulations, z_1 has no effect on w ($\pi_1 = 0$)[7].

Table 1 presents the average APEs estimates. For all three distributions of e, the two step estimation method provides more similar APEs estimates to the true APEs than the LCF approach. When w^2 is included in the models, allowing for the flexible forms of $\widehat{v}_i$ does not improve the approximation much. It provides almost the same estimates as allowing for only $\widehat{v}$.

Table 2 and Table 3 report the percentile APEs estimates with the normal and the χ^2_3 distributional assumptions, respectively. The results with the logistic distribution are similar to those with the normal distribution. In Table 2, the estimates by the two step estimation method under the normal distribution are similar to the true APEs across the percentiles of w distribution. Especially when w^2 is included, their biases are smaller than those by the LCF approach whether or not the flexible forms of $\widehat{v}_i$ is allowed. The estimation with the flexible forms of $\widehat{v}_i$ yields better estimates only for the 90^{th} percentile for the LCF approach. Table 3 also presents the similar results in general. But the approximation by the LCF approach is not good: its estimates at the 90th percentile of w distribution have the opposite directions to the true APEs.

Therefore, the simulations under Condition 1 suggest that the approximation by the two step estimation method even with a misspecified conditional mean is better than the LCF approach. Allowing for the additional terms of $\widehat{v}_i$ does not improve the approximation of the two approaches.

Condition 2

[7] I also conduct simulations by letting z_1 affect on w_i: $w_i = 0.5z_{i1} + z_{i2} + v_i$, $\tau = -0.5$ and $\sigma^2 = 1.25$. The results are similar as those under the Condition 1.

To examine if the results under Condition 1 are dependent on the instrument's predictive power, I change the values:

- $w_i = \pi_2 z_{i2} + v_i, \;\; \pi_2 \in \{0.1, 0.2, 0.5\}$
- $r_i = v_i + e_i$

Table 5 through Table 12 provide the simulation results under Condition 2 with the normal distributional and the χ^2_3 distributional assumptions[8]. The average APEs estimates with the different values of π_2 and the normal distributional assumption are shown in Table 5. According to the rule of the thumb suggested by Staiger and Stock (1997), the instrument with $\pi_2 = 0.1$ is considered as a weak instrument and the one with $\pi_2 \geq 0.2$ as a strong instrument. They discuss that the F statistic of the instruments in the first step dividing weak and strong instruments is 10 when there is one endogenous regressor. Table 4 illustrates the means of the first step's F statistics, which test the null hypothesis of π_2=0 in the simulations. The F statistics are larger than 10 when $\pi_2 \geq 0.2$.

Table 5 presents that both the two step estimation method and the LCF approach provide poor approximations with the weak instrument ($\pi_2 = 0.1$). But the approximation by the LCF approach is worse. When the nonlinearity of w is included, the LCF estimates are less biased than the two step method estimates. However, their standard errors are about ten times as large as the pseudo estimates'. Thus, their mean squared errors are much bigger than pseudo estimates' as shown in Table 13.

The two step estimation starts to recover its ability to approximate true APEs when the π_2 increases to 0.2, where the instrument becomes a strong one. Its estimates become less biased and less volatile than those with the weak instrument. However, the

[8]The logistic distributional assumption produces similar results with those under normal distributional assumption

estimates by the LCF approach still yields large biases although the standard deviations become smaller than those with the weak instrument. The LCF approach provides fairly good approximation when $\pi_2 = 0.5$. But the two step estimation still works better.

In addition, Table 5 shows that the additional CF terms in the estimation does not improve the approximation even when $\pi_2 < 1$. It causes a worse approximation for the two estimations.

The percentile APEs estimates under Condition 2 and the normally distributed e are reported in Table 6 through Table 8. They show similar results as Table 5 in general. First, Table 6 presents that the weak instrument causes the two estimation methods to produce the estimates with large biases and standard errors. Considering their mean squared errors in Table 14, the LCF approach provides worse estimates. Plus, the approximations by the two step estimation becomes better when π_2 increases to 0.2 even though those by the LCF approach does not. Table 7 shows that the pseudo estimates are less biased than those in Table 6 while the LCF estimates are not. The LCF approach manage to recover its approximation ability as π_2 rises to 0.5 as illustrated in Table 8. But still the two step estimation method's approximation is much better across the percentiles of w distribution: the pseudo APE estimates are much closer than the LCF estimates. Furthermore, as in the average APEs estimates, it is still hard to say that including the additional CF terms in the estimation improve the approximation. Especially, it brings explosive standard errors for the two step estimation with the weak instrument as shown in Table 6.

Table 9 through Table 12 are the results under Condition 2 and the χ^2_3 distributional assumption. The weak instrument also deteriorates the two approach's approximations like the two symmetric distributions. According to Table 9, it seems that the LCF approach can approximate the average APEs as well as the two step estimation method when $\pi_2 = 0.2$. But Table 11 present that only the two step estimation method provides the percentile estimates having a right direction across the the percentiles when $\pi_2 = 0.2$. Table 10 through Table 12 shows the smaller π_2 is, the more LCF percentile estimates with the opposite directions there are.

In conclusion, the simulation experiments provide evidence that although the quality of a instrument affects the two step estimation method's approximation, it is less sensitive to the weak instrument than the LCF approach's approximation.

4 Application: Michigan Educational Assessment Program math test

I apply the two step estimation method to Michigan Educational Assessment Program (MEAP) test of the school year 2004/2005 to estimate the effects of spending on students performance of the fourth grade math test outcomes. The fourth grade MEAP math test is a statewide assessment test given by the State Board of Education in Michigan. It measures fourth grade students' achievements in public schools in relation to Michigan curriculum standards that groups of educators, teachers and school administrators set. The students' outcomes are rated at one of the four levels as described in Table 15. Michigan Department of Education provides each district's percentage share of students for the four levels[9].

Papke (2005) and Papke and Wooldridge (2008) show that there exist nontrivial causal effects of spending on the pass rate of the fourth grade by using the MEAP math test data[10]. While they study the effects of spending on the pass rate with panel data set, this

[9] http://www.michigan.gov/mde/

[10] Papke (2005) uses school level data from 1992 to 1998 and Papke and Wooldridge (2008) uses district level data from 1992 to 2001. MEAP math test had 3 performance levels (Satisfactory, Moderate, Low)

application examines how spending shifts students in the four different levels instead of between pass and fail.

As in the two research, school spending in my application is also suspected to be endogenous. It is likely to be correlated with unobserved district effects such as parental involvement. For example, the parents who are enthusiastic and interested in their children's academic education are willing to provide extra learning opportunities for their children's achievements. And they are likely to put more pressure on schools to spend the resources more on students. Thus, it is necessary to use an instrumental variable to precisely estimate the effects of spending on students test outcomes. I use the same instrument, the foundation allowance that the two research use.

In 1994, Michigan reformed its school funding system with Proposal A. One of its objectives was to lower the school districts spending gap. It changed Michigan's school funding sources and imposed spending floors for school districts by providing minimum per pupil foundation allowances. It reduced the spending inequalities across the districts by allowing the low spending districts' foundation allowances to increase faster than other districts'. The initial foundation allowance for 1994/1995 was determined by a non-smooth function of per pupil spending in 1993/1994. For the following years, the incremental dollar increases are decided by comparing the previous year's foundation allowance with the basic foundation allowance that the legislature sets each year. There is no doubt that the foundation allowance is highly correlated with school spending and it is difficult to think it is correlated with the MEAP math test outcomes. Consequently, it becomes a natural instrument for spending.

The data set for the application contains 518 school districts[11]. I turn the percentage

before 2002. The pass rate, the dependent variable in the two research, measures the percent of students in the satisfactory level.

[11] The data does not contain charter schools (public school academies).

shares into proportions to obtain fractional dependent variables[12]. Table 16 illustrates the summary of the 518 districts' fractional dependent variables. While the lower corner 0 appears for all of the four levels, the upper corner 1 appears only for level 1 like the dependent variables generated in the simulations. Thus, I choose the first level as the reference choice.

In the estimation, I use log(per pupil expenditure) and log(foundation allowance) as spending and the instrument, respectively. I also control for the fraction of applications for the free and reduced-price lunch program as a measure of the poverty rate and log(enrollment) as a measure of school district size. Table 17 contains summary statistics of the explanatory variables in the data set.

Table 18 contains the first step estimation result. Netting out the other explanatory variables, the instrument's t statistic presents the strong correlation between the endogenous variable and the instrument. The F statistic also suggests that log(foundation allowance) is a strong instrument.

Table 19 reports the average APEs estimates. In general, the two estimation methods provide statistically significant estimates for level 1 and level 2. Without Including $\log(\text{spending})^2$ in the model, the estimates by the two step estimation method show that the conditional mean of level 1 is estimated to increase 2.4 percentage points and level 2's is estimated to decrease 1.7 percentage points if log(spending) increases by 0.1, which is about 10% increase in spending. The LCF estimates show larger effects. Considering that 119,687 students took the MEAP math exam in 2004/2005, one percentage point increase(decrease) in the number of students at a certain level represents about 1200 student increase(decrease) statewide. When the model includes $\log(\text{spending})^2$, the ef-

[12]The original percentage shares from MEAP may not sum to 100 because of rounding. Thus, I calculate the proportions not based on 100, but based on total percentage shares of the four levels.

fects become larger. The LCF approach has huge standard errors but allowing for $\hat{v}_i^2$ and $\hat{v}_i^3$ reduces them significantly.

Since the estimates of level 3 and level 4 are generally not statistically significant, average APEs estimates suggest that spending affects mainly on the top two levels. Plus, considering the magnitudes of the effects on level 1 and level 2 along with their directions, an increase in spending shifts the students who are rated at level 2 to level 1. Although the direct comparison is not valid because the performance level categories changed in 2002, the result of the application is consistent with the two research.

The percentile APEs estimates in Table 20 tell the similar story with the average APEs estimates. Interestingly, the estimates do not vary much across the percentiles. It seems that the spending variation in 2004/2005 is not big enough in the data.

5 Conclusion

This paper proposes a feasible two step estimation method for multiple fractional dependent variables with continuous endogenous explanatory variables. The method's important feature is that it suggests directly specifying the conditional mean as multinomial logit without the underlying assumptions. Monte Carlo simulation results provide evidence that the proposed estimation method is preferable to an alternate linear control function approach which does not account for the whole features of the fractional dependent variables even when the multinomial logit conditional mean is wrong. The application to fourth grade math test in Michigan illustrates that the proposed two step estimation method and the LCF approach provide similar results showing that an increase in spending shifts the students in the meeting Michigan standard to the exceeding Michigan standard.

References

6 Appendix

6.1 The adjusted standard errors of the two step estimation's APEs estimates

The adjusted standard errors of the parameter estimates by the two step estimation method should be calculated first in order to obtain the adjusted standard errors of its APEe estimates.

The first step in Procedure 2.1 obtains $\hat{v}_i$ from the regression w_i on z_i. Under the standard regularity conditions,

$$\sqrt{N}(\hat{\boldsymbol{\pi}} - \boldsymbol{\pi}) = N^{-\frac{1}{2}} \sum_{i}^{N} E(\mathbf{z}'\mathbf{z})^{-1}\mathbf{z}_i' v_i + o_p(1) = N^{-\frac{1}{2}} \sum_{i}^{N} \mathbf{q}_i + o_p(1) \tag{22}$$

where $\mathbf{q}_i \equiv E(\mathbf{z}_i'\mathbf{z}_i)^{-1}\mathbf{z}_i' v_i$ and as mentioned in the simulation setup, $\mathbf{z}_1$ contains a constant.

In the second step, I estimate $\boldsymbol{\theta}$ from the fmlogit $(y_{i1} \cdots y_{iG})$ on $(\mathbf{z}_{i1} w_i, \hat{v}_i)$ with (15) inserting $\hat{v}_i$ instead of v_i.

To calculate the adjusted standard errors, I redefine the parameter vectors as $\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\theta}_2' & \dots & \boldsymbol{\theta}_G' \end{pmatrix}'$ is a $p(G-1) \times 1$. In (15), $\boldsymbol{\theta}$ includes $\boldsymbol{\theta}_1$. But in this appendix, I drop the first choice parameters from $\boldsymbol{\theta}$ because they are defined as a zero vector.

The first order conditions is

$$\sum_i \mathbf{s}_i(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\pi}}) = \mathbf{0} \tag{23}$$

where

$$
\begin{aligned}
\mathbf{s}_i(\boldsymbol{\theta}) &\equiv \nabla_\theta \ell_i \\
&= \begin{pmatrix} (\frac{\partial \ell_i}{\partial \boldsymbol{\theta}_2})' \\ \vdots \\ (\frac{\partial \ell_i}{\partial \boldsymbol{\theta}_g})' \\ \vdots \\ (\frac{\partial \ell_i}{\partial \boldsymbol{\theta}_G})' \end{pmatrix} = \begin{pmatrix} \mathbf{s}_{i2} \\ \vdots \\ \mathbf{s}_{ig} \\ \vdots \\ \mathbf{s}_{iG} \end{pmatrix} \qquad s_{ig} = (\frac{\partial \ell_i}{\partial \boldsymbol{\theta}_g})'_{p\times 1} = \mathbf{x}_i'(y_{ig} - G_{ig})
\end{aligned}
$$

and $\mathbf{x} = \begin{pmatrix} \mathbf{z}_1 & w & v \end{pmatrix}$.

Using a mean value expansion (MVE) around $\boldsymbol{\theta}$, (23) is expressed as

$$
\sum_i \mathbf{s}_i(\widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\pi}}) = \sum_i \mathbf{s}_i(\boldsymbol{\theta}, \widehat{\boldsymbol{\pi}}) + \left[\nabla_\theta \sum_i \mathbf{s}_i(\ddot{\theta}) \right] (\widehat{\theta} - \theta) \tag{24}
$$

where $\ddot{\theta}$ is on the line segment between $\widehat{\theta}$ and θ. I rearrange (24) by multiplying by $\sqrt{N}$ and using (23).

$$
\sqrt{N}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) = E(-\mathbf{H}_i(\boldsymbol{\theta}))^{-1} \frac{1}{\sqrt{N}} \sum_i \mathbf{s}_i(\boldsymbol{\theta}, \widehat{\boldsymbol{\pi}}) + o_p(1) \tag{25}
$$

where

$$
\begin{aligned}
E[-\mathbf{H}_i] &\equiv \mathbf{A} \equiv E\left(-\nabla_\theta \mathbf{s}_i(\theta)\right) \\
&= E\begin{bmatrix}
\mathbf{x}_i'\mathbf{x}_i K_{i2}(1-K_{i2}) & -\mathbf{x}_i'\mathbf{x}_i K_{i2}K_{i3} & \cdots & \cdots & -\mathbf{x}_i'\mathbf{x}_i K_{i2}K_{iG} \\
-\mathbf{x}_i'\mathbf{x}_i K_{i3}K_{i2} & \mathbf{x}_i'\mathbf{x}_i K_{i3}(1-K_{i3}) & \cdots & \cdots & -\mathbf{x}_i'\mathbf{x}_i K_{i3}K_{iG} \\
\vdots & \vdots & & & \vdots \\
\vdots & \vdots & & & \vdots \\
-\mathbf{x}_i'\mathbf{x}_i K_{iG}K_{i2} & \cdots & \cdots & -\mathbf{x}_i'\mathbf{x}_i K_{iG}(K_{iG-1}) & \mathbf{x}_i'\mathbf{x}_i K_{iG}(1-K_{iG})
\end{bmatrix}
\end{aligned} \tag{26}
$$

Since $\sum_i \mathbf{s}_i(\boldsymbol{\theta}, \widehat{\boldsymbol{\pi}})$ depends on the sample, I can not apply a central limit theorem yet. By using a MVE around $\boldsymbol{\pi}$ again and multiplying by $\sqrt{N}$, I derive

$$
\frac{1}{\sqrt{N}}\sum_i \mathbf{s}_i(\boldsymbol{\theta}, \widehat{\boldsymbol{\pi}}) = \frac{1}{\sqrt{N}}\sum_i \mathbf{s}_i(\boldsymbol{\theta}, \boldsymbol{\pi}) + E[\nabla_\pi \mathbf{s}_i(\boldsymbol{\theta}, \boldsymbol{\pi})]\sqrt{N}(\widehat{\boldsymbol{\pi}} - \boldsymbol{\pi}) + o_p(1). \tag{27}
$$

By plugging (22) into (27), (27) is rewritten as

$$
\frac{1}{\sqrt{N}}\sum_i \mathbf{s}_i(\boldsymbol{\theta}, \widehat{\boldsymbol{\pi}}) = \frac{1}{\sqrt{N}}\sum_i \left(\mathbf{s}_i - \mathbf{F}\mathbf{q}_i\right) + o_p(1). \tag{28}
$$

where $\mathbf{F} = E[\nabla_\pi \mathbf{s}_i(\boldsymbol{\theta}, \boldsymbol{\pi})]_{p(G-1)\times M}$, $\nabla_\pi \mathbf{s}_i(\boldsymbol{\theta}, \boldsymbol{\pi}) = \frac{\partial \mathbf{x}_i'}{\partial \boldsymbol{\pi}}(y_{ig} - G_g) - \mathbf{x}_i' \frac{\partial G_g}{\partial \boldsymbol{\pi}}$.

By putting (28) into (25),

$$
\sqrt{N}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) = \mathbf{A}^{-1}\left[\frac{1}{\sqrt{N}}\sum_i \mathbf{d}_i(\boldsymbol{\theta}, \boldsymbol{\pi})\right] + o_p(1) \tag{29}
$$

where $\mathbf{d}_i \equiv \mathbf{s}_i - \mathbf{F}\mathbf{q}_i$.

Therefore, $Avar\left(\sqrt{N}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})\right) = \mathbf{A}^{-1}\mathbf{D}\mathbf{A}^{-1}$ where $\mathbf{D} \equiv Var(\mathbf{d}_i) = Var(\mathbf{s}_i - \mathbf{F}\mathbf{q}_i)$, and a valid estimator of $Avar\left(\sqrt{N}(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta})\right)$ is $\widehat{\mathbf{A}}^{-1}\widehat{\mathbf{D}}\widehat{\mathbf{A}}^{-1}$ where

$$\widehat{\mathbf{D}} \equiv \frac{1}{N}\sum_i \widehat{\mathbf{d}}_i\widehat{\mathbf{d}}_i' = \frac{1}{N}\sum_i (\widehat{\mathbf{s}}_i - \widehat{\mathbf{F}}\widehat{\mathbf{q}}_i)(\widehat{\mathbf{s}}_i - \widehat{\mathbf{F}}\widehat{\mathbf{q}}_i)', \tag{30}$$

$$\widehat{\mathbf{s}}_i = \mathbf{s}_i(\widehat{\mathbf{x}}_i, \widehat{\boldsymbol{\theta}}),$$

$$\widehat{\mathbf{F}}_i = \mathbf{F}_i(\widehat{\mathbf{x}}_i, \widehat{\boldsymbol{\theta}}),$$

$$\widehat{\mathbf{q}} \equiv (\frac{1}{N}\sum_i \mathbf{z}_i'\mathbf{z}_i)^{-1}\mathbf{z}_i'\widehat{v}_i, \quad \text{and}$$

$$\widehat{\mathbf{A}} \equiv \frac{1}{N}\sum_i -\widehat{\mathbf{H}}_i,$$

Consequently, the asymptotic variance of $\widehat{\boldsymbol{\theta}}$ is estimated as

$$\frac{1}{N}\widehat{\mathbf{A}}^{-1}\widehat{\mathbf{D}}\widehat{\mathbf{A}}^{-1}. \tag{31}$$

The two step estimation method's APEs estimate, evaluated at (z_{1k}, w_k), is

$$\widehat{\delta}^{[k]}_{wg} \equiv \frac{1}{N}\sum_i \widehat{G}^{[k]}_{ig}\left(\widehat{\theta}_{wg} - \frac{\sum_{h=2}\widehat{\theta}_{wh}\exp(\mathbf{z}_{1k}\widehat{\theta}_{zh} + w_k\widehat{\theta}_{wh} + \widehat{v}_i\widehat{\theta}_{vh})}{1+\sum_{h=2}\exp(\mathbf{z}_{1k}\widehat{\theta}_{zh} + w_k\widehat{\theta}_{wh} + \widehat{v}_i\widehat{\theta}_{vh})}\right) \tag{32}$$

where

$$\widehat{G}^{[k]}_{ig} = \widehat{G}_{ig}(\mathbf{z}_{1k}, w_k) = \frac{\exp(\mathbf{z}_{1k}\widehat{\theta}_{zg} + w_k\widehat{\theta}_{wg} + \widehat{v}_i\widehat{\theta}_{vg})}{1+\sum_{h=2}\exp(\mathbf{z}_{1k}\widehat{\theta}_{zh} + w_k\widehat{\theta}_{wh} + \widehat{v}_i\widehat{\theta}_{vh})} \quad g = 2, \cdots, G \quad \text{and}$$

$$\widehat{G}^{[k]}_{i1} = \widehat{G}_{i1}(\mathbf{z}_{1k}, w_k) = \frac{1}{1+\sum_{h=2}\exp(\mathbf{z}_{1k}\widehat{\theta}_{zh} + w_k\widehat{\theta}_{wh} + \widehat{v}_i\widehat{\theta}_{vh})} \quad g = 1.$$

Average APEs estimates are calculated as

$$
\begin{aligned}
\widehat{\delta}_{wg} &= \frac{1}{N}\sum_{k}\widehat{\delta}^{[k]}_{wg} \\
&= \frac{1}{N}\sum_{i}\left[\frac{1}{N}\sum_{k}\widehat{G}^{[k]}_{ig}\left(\widehat{\theta}_{wg} - \frac{\sum_{h=2}\widehat{\theta}_{wh}\exp(\mathbf{z}_{1k}\widehat{\theta}_{zh} + w_k\widehat{\theta}_{wh} + \widehat{v}_i\widehat{\theta}_{vh})}{1+\sum_{h=2}\exp(\mathbf{z}_{1k}\widehat{\theta}_{zh} + w_k\widehat{\theta}_{wh} + \widehat{v}_i\widehat{\theta}_{vh})}\right)\right] \\
&= \frac{1}{N}\sum_{i}\mathbf{j}_g(\mathbf{x}^{[k]}_i, \mathbf{z}_i, \widehat{\boldsymbol{\eta}})
\end{aligned}
\tag{33}
$$

where

$$
\mathbf{j}_g(\mathbf{x}^{[k]}_i, \mathbf{z}_i, \boldsymbol{\eta}) \equiv \frac{1}{N}\sum_{k}\left[G_{ig}[k]\left(\theta_{wg} - \frac{\sum_{h=2}\theta_{wh}\exp(x^{[k]}_i\theta_h)}{1+\sum_{h=2}\exp(x^{[k]}_i\theta_h)}\right)\right] \equiv \mathbf{j}_{ig}, \tag{34}
$$

$$
\boldsymbol{\eta} = \begin{pmatrix}\boldsymbol{\theta}\\ \\ \boldsymbol{\pi}\end{pmatrix}_{(p(G-1)+M)\times 1}, \quad \text{and}
$$

$$
x^{[k]}_i = (\mathbf{z}_{1k}, w_k, v_i).
$$

By using a MVE and subtracting $\sqrt{N}\delta_{wg}$ from both sides, I can derive

$$
\sqrt{N}(\widehat{\delta}_{wg} - \delta_{wg}) = \sqrt{N}\frac{1}{N}\sum_{i}\left(\mathbf{j}_{ig}(\mathbf{x}^{[k]}_i, \mathbf{z}_i, \boldsymbol{\eta}) - \delta_{wg}\right) + E\left[\nabla_{\eta}\mathbf{j}_{ig}\right]\sqrt{N}(\widehat{\boldsymbol{\eta}} - \boldsymbol{\eta}) + o_p(1). \tag{35}
$$

Using (29) and (22), (35) is expressed as

$$
\sqrt{N}(\widehat{\delta}_{wg} - \delta_{wg}) = \sqrt{N}\frac{1}{N}\sum_{i}\left(\mathbf{j}_{ig}(\mathbf{x}^{[k]}_i, \mathbf{z}_i, \boldsymbol{\eta}) - \delta_{wg} + E\left[\nabla_{\eta}\mathbf{j}_{ig}\right]\mathbf{k}_i\right) + o_p(1) \tag{36}
$$

where $\mathbf{k}_i = \begin{pmatrix} \mathbf{A}^{-1}\mathbf{d}_i \\ \mathbf{q}_i \end{pmatrix}$. Since $E\left(\mathbf{j}_{ig}(\mathbf{x}_i^{[k]}, \mathbf{z}_i, \boldsymbol{\eta}) - \delta_{wg} + E\left[\nabla_\eta \mathbf{j}_{ig}\right]\mathbf{k}_i\right) = 0$,

$$Avar(\sqrt{N}(\widehat{\delta}_{wg} - \delta_{wg})) = Var\left(\mathbf{j}_{ig}(\mathbf{x}_i^{[k]}, \mathbf{z}_i, \boldsymbol{\eta}) - \delta_{wg} + \mathbf{J}_g(\eta)\mathbf{k}_i\right) \tag{37}$$

where $\mathbf{J}_g(\eta) \equiv E\left[\nabla_\eta \mathbf{j}_{ig}\right]$, a $1 \times p$ jacobian, and a valid estimator of $Avar\left(\sqrt{N}(\widehat{\delta}_{wg} - \delta_{wg})\right)$ is

$$\frac{1}{N}\sum_i \left[\widehat{\mathbf{j}}_{ig} - \widehat{\delta}_{wg} + \widehat{\mathbf{J}}_g\widehat{\mathbf{k}}_i\right]\left[\widehat{\mathbf{j}}_{ig} - \widehat{\delta}_{wg} + \widehat{\mathbf{J}}_g\widehat{\mathbf{k}}_i\right]' \tag{38}$$

where

$$\begin{aligned}
\widehat{\mathbf{j}}_{ig} &= \left[\frac{1}{N}\sum_k \widehat{K}_{ig}[k]\left(\widehat{\theta}_{wg} - \frac{\sum_{h=2}\widehat{\theta}_{wh}\exp(\widehat{\mathbf{x}}_i^{[k]}\widehat{\boldsymbol{\theta}}_h)}{1 + \sum_{h=2}\exp(\widehat{\mathbf{x}}_i^{[k]}\widehat{\boldsymbol{\theta}}_h)}\right)\right], \\
\widehat{\delta}_{wg} &= \frac{1}{N}\sum_i \widehat{\mathbf{j}}_{ig}, \\
\widehat{\mathbf{k}}_i &= \begin{pmatrix} \widehat{\mathbf{A}}^{-1}\widehat{\mathbf{d}}_i \\ \widehat{\mathbf{q}}_i \end{pmatrix}, \quad \text{and} \\
\widehat{\mathbf{J}}_g &= \frac{1}{N}\left[\nabla_\eta \mathbf{j}_{ig}(\mathbf{x}_i^{[k]}, \widehat{\eta})\right].
\end{aligned}$$

The adjusted standard errors of $\widehat{\delta}_{wg}$ is obtained as the square root of (38) multiplied by $\frac{1}{\sqrt{N}}$.

For the percentile APEs estimates' standard errors, the same process is used by defining $\widehat{\delta}_{wg}$ and $\widehat{\mathbf{j}}_{ig}$ as the following:

$$\widehat{\delta}_{wg} = \frac{1}{N}\sum_i \widehat{G}_{ig}(\overline{\mathbf{z}}_1, w_p)\left(\widehat{\theta}_{wg} - \frac{\sum_{h=2}\widehat{\theta}_{wh}\exp(\overline{\mathbf{z}}_1\widehat{\theta}_{zh} + w_p\widehat{\theta}_{wh} + \widehat{v}_i\widehat{\theta}_{vh})}{1+\sum_{h=2}\exp(\overline{\mathbf{z}}_1\widehat{\theta}_{zh} + w_p\widehat{\theta}_{wh} + \widehat{v}_i\widehat{\theta}_{vh})}\right),$$
$$\widehat{\mathbf{j}}_{ig} = \widehat{G}_{ig}(\overline{\mathbf{z}}_1 w_p)\left(\widehat{\theta}_{wg} - \frac{\sum_{h=2}\widehat{\theta}_{wh}\exp(\overline{\mathbf{z}}_1\widehat{\theta}_{zh} + w_p\widehat{\theta}_{wh} + \widehat{v}_i\widehat{\theta}_{vh})}{1+\sum_{h=2}\exp(\overline{\mathbf{z}}_1\widehat{\theta}_{zh} + w_p\widehat{\theta}_{wh} + \widehat{v}_i\widehat{\theta}_{vh})}\right).$$

When the model includes w^2 and the flexible forms of $\widehat{v}$ is used, the process with including them can provide the adjusted standard errors.

6.2 The adjusted standard errors of LCF estimates

I calculate the adjusted standard errors of the LCF estimates following Wooldridge (2010, Appendix 6A). For $g = 2, \cdots, G$, a researcher using the LCF approach models the following:

$$y_{ig} = \mathbf{z}_{i1}\boldsymbol{\gamma}_{zg} + \gamma_{wg}w_i + u_{ig} \tag{39}$$

The reduced form for the endogenous variable and the errors are written as, respectively,

$$w_i = \mathbf{z}_i\boldsymbol{\pi} + v_i \tag{40}$$
$$u_{ig} = \rho_g v_i + e_{ig} \tag{41}$$

where the dimensions of the variables and the parameters are the same as before. Then using (41), (39) is rewritten

$$y_{ig} = \mathbf{z}_{i1}\boldsymbol{\gamma}_{zg} + \gamma_{wg}w_i + \rho_g v_i + e_{ig} = \mathbf{x}_i\boldsymbol{\gamma}_g + e_{ig} \tag{42}$$

where $\boldsymbol{\gamma}_g = \begin{pmatrix} \boldsymbol{\gamma}_{zg} \\ \gamma_{wg} \\ \rho_{vg} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\gamma}_{zg} \\ \gamma_{wg} \\ \gamma_{vg} \end{pmatrix}$.

The second step of Procedure 3.1 uses (42) replacing v with $\widehat{v}$ from the first step. The estimating equation is

$$y_{ig} = \mathbf{z}_{i1}\boldsymbol{\gamma}_{zg} + \gamma_{wg} w_i + \rho_g \widehat{v}_i + \left(\rho_g(v_i - \widehat{v}_i) + e_{ig}\right) = \widehat{\mathbf{x}}_{vi}\boldsymbol{\gamma}_g + \left(\rho_g(v_i - \widehat{v}_i) + e_{ig}\right) \tag{43}$$

$$= \widehat{\mathbf{x}}_{vi}\boldsymbol{\gamma}_g + (\mathbf{x}_i - \widehat{\mathbf{x}}_i)\boldsymbol{\gamma}_g + e_{ig}. \tag{44}$$

The OLS estimator is expressed as

$$\widehat{\boldsymbol{\gamma}}_g = (\sum_i \widehat{\mathbf{x}}'_{vi}\widehat{\mathbf{x}}_{vi})^{-1}\sum_i \widehat{\mathbf{x}}'_{vi} y_{ig} = \boldsymbol{\gamma}_g + (\sum_i \widehat{\mathbf{x}}'_{vi}\widehat{\mathbf{x}}_{vi})^{-1}\sum_i \widehat{\mathbf{x}}'_{vi}\left((\mathbf{x}_i - \widehat{\mathbf{x}}_{vi})\boldsymbol{\gamma}_g + e_{ig}\right), \tag{45}$$

and I can derive

$$\sqrt{N}(\widehat{\boldsymbol{\gamma}}_g - \boldsymbol{\gamma}_g) = (\frac{1}{N}\sum_i \widehat{\mathbf{x}}'_i\widehat{\mathbf{x}}_i)^{-1}\frac{1}{\sqrt{N}}\sum_i \widehat{\mathbf{x}}'_i\left((\mathbf{x}_i - \widehat{\mathbf{x}}_i)\boldsymbol{\gamma}_g + e_{ig}\right) \tag{46}$$

Using MVEs, $vec(AXB) = (B' \otimes A)vec(X)$ and a weak law of large numbers, (46) can be expressed as

$$\sqrt{N}\left(\widehat{\boldsymbol{\gamma}}_g - \boldsymbol{\gamma}_g\right) = \mathbf{C}^{-1}\left\{\frac{1}{\sqrt{N}}\sum_i^N \left(\mathbf{x}'_i e_{ig} - \mathbf{R}_g\mathbf{A}^{-1}\mathbf{z}'_i v_i\right)\right\} + o_p(1) \tag{47}$$

where

$$\mathbf{C} \equiv E(\mathbf{x}'\mathbf{x}),$$

$$\mathbf{R}_g = E\left[(\boldsymbol{\gamma}_g \otimes \mathbf{x})'\nabla_\pi \mathbf{x}(\boldsymbol{\pi})\right], \text{ and}$$

$$\mathbf{A} = E(\mathbf{z}'\mathbf{z}).$$

By the central limit theorem,

$$\sqrt{N}\left(\widehat{\boldsymbol{\gamma}}_g - \boldsymbol{\gamma}_g\right) \overset{a}{\sim} Normal\left[0, \mathbf{C}^{-1}\mathbf{M}_g\mathbf{C}^{-1}\right] \tag{48}$$

where $\mathbf{M}_g = Var\left(\mathbf{x}'e_g - \mathbf{R}_g\mathbf{A}^{-1}\mathbf{z}'v\right)$. Asymptotic variance of $\widehat{\boldsymbol{\gamma}}_g$ is estimated as

$$\widehat{\mathbf{C}}^{-1}\widehat{\mathbf{M}}_g\widehat{\mathbf{C}}^{-1}/N \tag{49}$$

where

$$\widehat{\mathbf{C}} = \frac{1}{N}\sum_i^N \widehat{\mathbf{x}}'_{vi}\widehat{\mathbf{x}}_{vi},$$

$$\widehat{\mathbf{M}}_g = \frac{1}{N}\sum_i^N \left(\widehat{\mathbf{x}}'_{vi}\widehat{e}_{ig} - \widehat{\mathbf{G}}_g\widehat{\mathbf{A}}^{-1}\mathbf{z}'_i\widehat{v}_i\right)\left(\widehat{\mathbf{x}}'_{vi}\widehat{e}_{ig} - \widehat{\mathbf{G}}_g\widehat{\mathbf{A}}^{-1}\mathbf{z}'_i\widehat{v}_i\right)' \quad \text{for } g = 2, \cdots, G,$$

$$\widehat{\mathbf{R}}_g = \frac{1}{N}\sum_i^N \left(\widehat{\boldsymbol{\gamma}}_g \otimes \widehat{\mathbf{x}}_{vi}\right)'\nabla_\pi \mathbf{x}_i(\widehat{\boldsymbol{\pi}}) \quad \text{for } g = 2, \cdots, G,$$

$$\widehat{\mathbf{A}} = \frac{1}{N}\sum_i^N \mathbf{z}'_i\mathbf{z}_i, \text{ and}$$

$$\widehat{e}_{ig} = y_{ig} - \widehat{\mathbf{x}}_i\widehat{\boldsymbol{\gamma}}_g \quad \text{for } g = 2, \cdots, G.$$

The first choice estimate is $\widehat{\boldsymbol{\gamma}}_1 = \mathbf{e}_1 - \sum_{g=2}^G \widehat{\boldsymbol{\gamma}}_g$ where $\mathbf{e}_1 = \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix}'_{G\times 1}$. Using

(43) and multiplying by $\sqrt{N}$, $\sqrt{N}(\widehat{\boldsymbol{\gamma}}_1 - \boldsymbol{\gamma}_1)$ is written as

$$\sqrt{N}\left(\widehat{\boldsymbol{\gamma}}_1 - \boldsymbol{\gamma}_1\right) = \left(\frac{1}{N}\sum_i \widehat{\mathbf{x}}'_{vi}\widehat{\mathbf{x}}_{vi}\right)^{-1}\left(-\frac{1}{\sqrt{N}}\sum_i \widehat{\mathbf{x}}'_{vi}\left[\sum_{g=2}^{G}\left((\mathbf{x}_i - \widehat{\mathbf{x}}_{vi})\widehat{\boldsymbol{\gamma}}_g + e_{ig}\right)\right]\right). \tag{50}$$

Using (46) and (47), (50) becomes

$$\sqrt{N}(\widehat{\boldsymbol{\gamma}}_1 - \boldsymbol{\gamma}_1) = \mathbf{C}^{-1}\left(-\frac{1}{\sqrt{N}}\sum_i^N\sum_{g=2}^G\left(\mathbf{x}'_i e_{ig} - \mathbf{R}_g\mathbf{A}^{-1}\mathbf{z}'_i v_i\right)\right) + o_p(1) \tag{51}$$

$$\overset{a}{\sim} Normal\left[0, \mathbf{C}^{-1}\mathbf{M}_1\mathbf{C}^{-1}\right] \tag{52}$$

where $\mathbf{M}_1 = Var\left(\sum_{g=2}^G\left(\mathbf{x}'_i e_{ig} - \mathbf{R}_g\mathbf{A}^{-1}\mathbf{z}'_i v_i\right)\right)$. Then a valid estimator of $Avar(\widehat{\boldsymbol{\gamma}}_1)$ is

$$\widehat{\mathbf{C}}^{-1}\widehat{\mathbf{M}}_1\widehat{\mathbf{C}}^{-1}/N \tag{53}$$

where

$$\widehat{\mathbf{M}}_1 = \sum_{g=2}^G\widehat{\mathbf{M}}_g + \frac{1}{N}\sum_{g\neq k}\left[\left(\mathbf{x}'_i\widehat{e}_{ig} - \mathbf{R}_g\mathbf{A}^{-1}\mathbf{z}'_i v_i\right)\left(\mathbf{x}'_i\widehat{e}_{ik} - \mathbf{R}_k\mathbf{A}^{-1}\mathbf{z}'_i v_i\right)'\right].$$

The adjusted standard errors of $\widehat{\boldsymbol{\gamma}}_g$ is obtained by the square roots of the diagonal elements of (49) and (53). In addition, as I mentioned at the end of 7.1, the adjusted standard errors of the additional cases need slight modifications.

7 Tables

Table 1: Average APEs under Condition 1

$D(e)$			Normal			Logistic			χ^2_3		
g			1	2	3	1	2	3	1	2	3
w	True	Mean	-0.117	0.059	0.059	-0.108	0.054	0.054	-0.049	0.025	0.025
	Two Step	Mean	-0.117	0.058	0.058	-0.108	0.054	0.054	-0.047	0.024	0.024
		SD	0.009	0.005	0.005	0.011	0.006	0.006	0.007	0.004	0.004
		SE	-	-	-	-	-	-	-	-	-
	LCF	Mean	-0.109	0.054	0.054	-0.101	0.051	0.051	-0.052	0.026	0.026
		SD	0.009	0.005	0.005	0.011	0.006	0.006	0.008	0.005	0.005
		SE	0.009	0.005	0.005	0.011	0.006	0.006	0.008	0.005	0.005
w^2	True	Mean	-0.127	0.063	0.063	-0.116	0.058	0.058	-0.061	0.030	0.030
	Two Step	Mean	-0.126	0.063	0.063	-0.115	0.058	0.057	-0.059	0.030	0.030
		SD	0.009	0.005	0.005	0.012	0.006	0.006	0.008	0.004	0.005
		SE	-	-	-	-	-	-	-	-	-
	Two Step	Mean	-0.126	0.063	0.063	-0.115	0.057	0.057	-0.058	0.029	0.029
	(Flexible)	SD	0.009	0.005	0.005	0.012	0.006	0.006	0.008	0.004	0.004
		SE	-	-	-	-	-	-	-	-	-
	LCF	Mean	-0.117	0.059	0.059	-0.108	0.054	0.054	-0.064	0.032	0.032
		SD	0.009	0.005	0.005	0.011	0.006	0.006	0.008	0.005	0.005
		SE	0.009	0.005	0.005	0.011	0.006	0.006	0.008	0.004	0.004
	LCF	Mean	-0.118	0.059	0.059	-0.108	0.054	0.054	-0.064	0.032	0.032
	(Flexible)	SD	0.009	0.005	0.005	0.011	0.006	0.006	0.008	0.005	0.005
		SE	0.008	0.005	0.005	0.011	0.006	0.006	0.008	0.004	0.004

[1] $\pi_1 = 0, \pi_2 = 1, \rho = 1$

Table 2: Percentile APEs under Condition 1 and Normal distribution

w_p			10^{th}			25^{th}			50^{th}			75^{th}			90^{th}		
g			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
w	True	Mean	-0.181	0.090	0.090	-0.171	0.085	0.085	-0.132	0.066	0.066	-0.085	0.043	0.043	-0.050	0.025	0.025
	Two Step	Mean	-0.184	0.092	0.092	-0.173	0.087	0.087	-0.131	0.065	0.065	-0.082	0.041	0.041	-0.048	0.024	0.024
		SD	0.017	0.009	0.009	0.019	0.010	0.010	0.012	0.007	0.006	0.006	0.004	0.004	0.006	0.004	0.004
		SE	0.016	0.008	0.008	0.018	0.009	0.009	0.011	0.006	0.006	0.004	0.003	0.003	0.003	0.003	0.003
w^2	True	Mean	-0.241	0.121	0.121	-0.206	0.103	0.103	-0.132	0.066	0.066	-0.073	0.036	0.036	-0.040	0.020	0.020
	Two Step	Mean	-0.240	0.120	0.120	-0.208	0.104	0.104	-0.132	0.066	0.066	-0.072	0.036	0.036	-0.039	0.020	0.019
		SD	0.023	0.011	0.012	0.022	0.012	0.012	0.014	0.008	0.007	0.009	0.005	0.005	0.007	0.005	0.005
		SE	0.031	0.010	0.010	0.023	0.008	0.008	0.012	0.005	0.005	0.007	0.005	0.005	0.006	0.008	0.008
	Two Step	Mean	-0.239	0.120	0.120	-0.208	0.104	0.104	-0.132	0.066	0.066	-0.071	0.036	0.035	-0.038	0.019	0.019
	(Flexible)	SD	0.023	0.012	0.012	0.022	0.012	0.012	0.014	0.008	0.008	0.009	0.005	0.005	0.007	0.006	0.006
		SE	0.031	0.011	0.011	0.023	0.009	0.009	0.012	0.005	0.005	0.008	0.005	0.005	0.007	0.009	0.009
	LCF	Mean	-0.206	0.103	0.103	-0.164	0.082	0.082	-0.118	0.059	0.059	-0.071	0.036	0.035	-0.029	0.014	0.014
		SD	0.018	0.009	0.009	0.013	0.007	0.007	0.009	0.005	0.005	0.010	0.006	0.006	0.015	0.008	0.009
		SE	0.016	0.008	0.008	0.012	0.006	0.006	0.009	0.005	0.005	0.010	0.006	0.006	0.014	0.008	0.008
	LCF	Mean	-0.190	0.095	0.095	-0.156	0.078	0.078	-0.118	0.059	0.059	-0.079	0.040	0.040	-0.045	0.023	0.022
	(Flexible)	SD	0.018	0.009	0.009	0.013	0.007	0.007	0.009	0.005	0.005	0.009	0.005	0.005	0.014	0.008	0.008
		SE	0.016	0.009	0.009	0.012	0.006	0.006	0.008	0.005	0.005	0.009	0.005	0.005	0.012	0.007	0.007

1. APEs at $(\overline{z}_1, w_p)$.

2. $\pi_1 = 0, \pi_2 = 1, \rho = 1$.

Table 3: Percentile APEs under Condition 1 and χ_3^2 distribution

w_p			10^{th}			25^{th}			50^{th}			75^{th}			90^{th}		
g			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
w	True	mean	-0.100	0.050	0.050	-0.068	0.034	0.034	-0.038	0.019	0.019	-0.018	0.009	0.009	-0.009	0.004	0.004
	Two Step	Mean	-0.096	0.048	0.048	-0.061	0.030	0.030	-0.033	0.017	0.017	-0.017	0.009	0.009	-0.009	0.005	0.005
		SD	0.022	0.011	0.011	0.011	0.006	0.006	0.004	0.003	0.003	0.002	0.003	0.003	0.002	0.003	0.003
		SE	0.021	0.010	0.010	0.010	0.005	0.005	0.003	0.003	0.003	0.001	0.003	0.003	0.001	0.003	0.003
w^2	True	Mean	-0.152	0.076	0.076	-0.086	0.043	0.043	-0.038	0.019	0.019	-0.016	0.008	0.008	-0.007	0.004	0.004
	Two Step	Mean	-0.134	0.067	0.067	-0.080	0.040	0.040	-0.040	0.020	0.020	-0.018	0.009	0.009	-0.008	0.004	0.004
		SD	0.032	0.016	0.016	0.014	0.007	0.007	0.007	0.004	0.004	0.003	0.003	0.003	0.002	0.004	0.004
		SE	0.030	0.020	0.020	0.012	0.012	0.012	0.006	0.006	0.006	0.003	0.003	0.003	0.002	0.013	0.013
	Two Step (Flexible)	Mean	-0.133	0.066	0.066	-0.078	0.039	0.039	-0.038	0.019	0.019	-0.017	0.009	0.009	-0.008	0.004	0.004
		SD	0.031	0.016	0.016	0.013	0.007	0.007	0.007	0.004	0.004	0.004	0.004	0.003	0.002	0.004	0.004
		SE	0.030	0.023	0.023	0.012	0.015	0.015	0.007	0.008	0.008	0.004	0.005	0.005	0.002	0.016	0.016
	LCF	Mean	-0.162	0.081	0.081	-0.116	0.058	0.058	-0.064	0.032	0.032	-0.013	0.006	0.006	0.034	-0.017	-0.017
		SD	0.019	0.010	0.010	0.013	0.007	0.007	0.009	0.005	0.005	0.008	0.005	0.005	0.011	0.007	0.007
		SE	0.017	0.009	0.009	0.012	0.007	0.007	0.008	0.005	0.005	0.008	0.005	0.005	0.011	0.006	0.006
	LCF (Flexible)	Mean	-0.143	0.071	0.071	-0.105	0.053	0.053	-0.064	0.032	0.032	-0.023	0.011	0.011	0.015	-0.007	-0.007
		SD	0.022	0.011	0.011	0.015	0.008	0.008	0.009	0.005	0.005	0.007	0.004	0.004	0.011	0.007	0.007
		SE	0.020	0.010	0.010	0.014	0.007	0.007	0.008	0.004	0.004	0.006	0.004	0.004	0.010	0.006	0.006

1. APEs at $(\bar{z}_1, w_p)$.
2. $\pi_1 = 0, \pi_2 = 1, \rho = 1$.
3. The grey colored cells indicate that the estimates have the opposite directions to the true APEs.

Table 4: F statistics of the first step

π_2		0.1	0.2	0.5	1
F statistic	Mean	3.682	11.491	72.848	499.975
	SD	3.618	6.781	18.436	63.747

1. $H_0 : \pi_2 = 0$.

Table 5: Average APEs under Condition 2 and Normal distribution

π_2			0.1			0.2			0.5			1		
g			1	2	3	1	2	3	1	2	3	1	2	3
w	True	Mean	-0.113	0.056	0.056	-0.113	0.056	0.056	-0.114	0.057	0.057	-0.117	0.059	0.059
	Two Step	Mean	-0.094	0.046	0.047	-0.100	0.050	0.050	-0.112	0.056	0.056	-0.117	0.058	0.058
		SD	0.103	0.062	0.056	0.061	0.033	0.032	0.020	0.011	0.011	0.009	0.005	0.005
		SE	-	-	-	-	-	-	-	-	-	-	-	-
	LCF	Mean	-0.079	0.042	0.037	-0.071	0.036	0.036	-0.100	0.050	0.050	-0.109	0.054	0.054
		SD	1.101	0.731	0.427	0.390	0.188	0.203	0.021	0.011	0.011	0.009	0.005	0.005
		SE	6.501	4.264	2.297	0.568	0.275	0.296	0.020	0.011	0.011	0.009	0.005	0.005
w^2	True	Mean	-0.121	0.060	0.060	-0.121	0.061	0.061	-0.122	0.061	0.061	-0.127	0.063	0.063
	Two Step	Mean	-0.100	0.050	0.050	-0.106	0.053	0.054	-0.120	0.060	0.060	-0.126	0.063	0.063
		SD	0.106	0.060	0.060	0.064	0.035	0.032	0.021	0.011	0.011	0.009	0.005	0.005
		SE	-	-	-	-	-	-	-	-	-	-	-	-
	Two Step (Flexible)	Mean	-0.094	0.046	0.047	-0.102	0.051	0.052	-0.119	0.059	0.059	-0.126	0.063	0.063
		SD	0.104	0.059	0.060	0.063	0.035	0.033	0.021	0.012	0.011	0.009	0.005	0.005
		SE	-	-	-	-	-	-	-	-	-	-	-	-
	LCF	Mean	-0.114	0.059	0.055	-0.080	0.039	0.041	-0.108	0.054	0.054	-0.117	0.059	0.059
		SD	1.262	0.584	0.705	0.397	0.227	0.171	0.020	0.011	0.010	0.009	0.005	0.005
		SE	7.251	3.272	4.003	0.577	0.327	0.254	0.019	0.010	0.010	0.009	0.005	0.005
	LCF (Flexible)	Mean	-0.110	0.057	0.053	-0.077	0.037	0.040	-0.108	0.054	0.054	-0.118	0.059	0.059
		SD	1.381	0.651	0.755	0.467	0.268	0.200	0.019	0.011	0.010	0.009	0.005	0.005
		SE	7.877	3.547	4.350	0.670	0.382	0.292	0.019	0.010	0.010	0.008	0.005	0.005

1. $\pi_1 = 0$, $\rho = 1$.

Table 6: Percentile APEs under Condition 2, $\pi_2 = 0.1$, and Normal distribution

w_p			10^{th}			25^{th}			50^{th}			75^{th}			90^{th}		
g			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
w	True	Mean	-0.164	0.082	0.082	-0.156	0.078	0.078	-0.127	0.064	0.064	-0.089	0.044	0.044	-0.057	0.028	0.028
	Two Step	Mean	-0.133	0.065	0.068	-0.170	0.084	0.086	-0.120	0.059	0.061	-0.049	0.025	0.024	-0.017	0.009	0.008
		SD	0.117	0.066	0.060	0.168	0.092	0.086	0.138	0.082	0.073	0.084	0.058	0.052	0.059	0.045	0.042
		SE	0.186	0.112	0.104	0.215	0.122	0.113	0.198	0.114	0.118	0.136	0.098	0.089	0.136	0.092	0.103
w^2	True	Mean	-0.219	0.110	0.110	-0.188	0.094	0.094	-0.127	0.064	0.064	-0.075	0.037	0.037	-0.044	0.022	0.022
	Two Step	Mean	-0.160	0.080	0.080	-0.198	0.099	0.099	-0.122	0.061	0.062	-0.041	0.020	0.021	-0.007	0.003	0.004
		SD	0.121	0.065	0.066	0.178	0.093	0.094	0.139	0.077	0.079	0.085	0.054	0.055	0.061	0.042	0.043
		SE	0.256	3.819	10.748	0.243	2.116	5.910	0.191	1.227	3.343	0.147	0.292	0.245	0.192	2.059	6.471
	Two Step (Flexible)	Mean	-0.155	0.076	0.078	-0.193	0.095	0.098	-0.122	0.060	0.061	-0.036	0.018	0.018	0.009	-0.004	-0.005
		SD	0.141	0.080	0.079	0.178	0.097	0.097	0.146	0.080	0.083	0.106	0.063	0.066	0.097	0.061	0.061
		SE	2E+17	8E+28	6E+29	9E+04	1E+29	5E+29	6E+09	8E+28	5E+29	8E+19	9E+28	5E+29	1E+25	2E+29	2E+29
		SE*	0.222	0.155	0.154	0.190	0.159	0.155	0.121	0.095	0.093	0.088	0.051	0.059	0.104	0.113	0.122
	LCF	Mean	-0.213	0.109	0.104	-0.166	0.085	0.081	-0.114	0.059	0.055	-0.062	0.033	0.029	-0.015	0.010	0.005
		SD	1.262	0.584	0.705	1.262	0.584	0.705	1.262	0.584	0.705	1.262	0.584	0.704	1.261	0.584	0.704
		SE	7.253	3.273	4.004	7.252	3.272	4.004	7.251	3.272	4.003	7.252	3.272	4.003	7.253	3.273	4.004
	LCF (Flexible)	Mean	-0.178	0.092	0.086	-0.146	0.076	0.070	-0.110	0.057	0.053	-0.075	0.039	0.036	-0.042	0.022	0.020
		SD	1.477	0.692	0.808	1.407	0.660	0.771	1.378	0.649	0.754	1.422	0.676	0.771	1.507	0.722	0.810
		SE	8.089	3.649	4.461	7.955	3.585	4.389	7.877	3.547	4.350	7.918	3.566	4.380	8.033	3.620	4.446

1. APEs at $(\overline{z}_1, w_p)$.
2. $\pi_1 = 0$, $\rho = 1$.
3. SE* is the median of the standard errors.
4. The grey colored cells indicate that the estimates have the opposite directions to the true APEs.

Table 7: Percentile APEs under Condition 2, $\pi_2 = 0.2$, and Normal distribution

w_p			10^{th}			25^{th}			50^{th}			75^{th}			90^{th}		
g			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
w	True	Mean	-0.164	0.082	0.082	-0.157	0.078	0.078	-0.127	0.064	0.064	-0.089	0.044	0.044	-0.057	0.028	0.028
	Two Step	Mean	-0.151	0.075	0.076	-0.156	0.077	0.078	-0.115	0.057	0.058	-0.067	0.033	0.034	-0.037	0.019	0.019
		SD	0.087	0.045	0.044	0.103	0.053	0.052	0.076	0.041	0.039	0.046	0.028	0.027	0.033	0.022	0.021
		SE	0.084	0.043	0.043	0.100	0.053	0.052	0.076	0.039	0.042	0.040	0.026	0.029	0.034	0.026	0.024
w^2	True	Mean	-0.220	0.110	0.110	-0.189	0.094	0.094	-0.127	0.064	0.064	-0.075	0.037	0.037	-0.043	0.022	0.022
	Two Step	Mean	-0.190	0.094	0.095	-0.184	0.092	0.093	-0.117	0.058	0.059	-0.057	0.028	0.029	-0.025	0.012	0.013
		SD	0.090	0.047	0.045	0.110	0.058	0.055	0.076	0.043	0.039	0.047	0.029	0.026	0.037	0.024	0.023
		SE	0.137	4.547	0.109	0.120	1.900	0.075	0.072	0.510	0.034	0.042	0.033	0.027	0.045	0.072	0.033
	Two Step (Flexible)	Mean	-0.187	0.093	0.094	-0.183	0.091	0.093	-0.116	0.057	0.059	-0.053	0.026	0.027	-0.015	0.008	0.007
		SD	0.102	0.054	0.054	0.111	0.059	0.057	0.078	0.044	0.041	0.060	0.034	0.033	0.062	0.036	0.037
		SE	0.314	14.131	0.578	0.200	5.955	0.312	0.120	1.580	0.105	0.099	0.093	0.044	0.294	0.392	0.122
	LCF	Mean	-0.179	0.088	0.091	-0.132	0.065	0.067	-0.080	0.039	0.041	-0.028	0.013	0.015	0.019	-0.011	-0.008
		SD	0.397	0.227	0.171	0.397	0.227	0.171	0.397	0.227	0.171	0.398	0.227	0.172	0.398	0.227	0.172
		SE	0.580	0.328	0.255	0.578	0.327	0.254	0.577	0.327	0.254	0.577	0.327	0.254	0.579	0.328	0.255
	LCF (Flexible)	Mean	-0.138	0.068	0.070	-0.109	0.053	0.056	-0.077	0.037	0.040	-0.045	0.021	0.024	-0.016	0.006	0.010
		SD	0.499	0.285	0.215	0.481	0.275	0.206	0.467	0.268	0.200	0.459	0.264	0.197	0.458	0.264	0.197
		SE	0.698	0.395	0.307	0.681	0.387	0.298	0.669	0.382	0.292	0.670	0.385	0.293	0.683	0.395	0.300

1. APEs at $(\overline{z}_1, w_p)$.

2. $\pi_1 = 0$, $\rho = 1$.

3. The grey colored cells indicate that the estimates have the opposite directions to the true APEs.

Table 8: Percentile APEs under Condition 2, $\pi_2 = 0.5$, and Normal distribution

w_p			10^{th}			25^{th}			50^{th}			75^{th}			90^{th}		
g			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
w	True	Mean	-0.167	0.084	0.084	-0.159	0.080	0.080	-0.128	0.064	0.064	-0.088	0.044	0.044	-0.055	0.028	0.028
	Two Step	Mean	-0.168	0.084	0.084	-0.162	0.081	0.081	-0.127	0.063	0.063	-0.083	0.042	0.042	-0.051	0.026	0.026
		SD	0.038	0.019	0.019	0.041	0.021	0.020	0.026	0.014	0.014	0.011	0.008	0.007	0.007	0.007	0.007
		SE	0.036	0.018	0.018	0.039	0.020	0.020	0.025	0.013	0.013	0.008	0.007	0.007	0.006	0.007	0.007
w^2	True	Mean	-0.224	0.112	0.112	-0.192	0.096	0.096	-0.128	0.064	0.064	-0.074	0.037	0.037	-0.043	0.021	0.021
	Two Step	Mean	-0.216	0.108	0.108	-0.192	0.096	0.096	-0.129	0.064	0.064	-0.073	0.036	0.037	-0.041	0.020	0.020
		SD	0.041	0.020	0.020	0.045	0.023	0.022	0.027	0.015	0.014	0.014	0.009	0.009	0.010	0.008	0.008
		SE	0.060	0.013	0.013	0.048	0.012	0.012	0.024	0.010	0.010	0.012	0.008	0.008	0.011	0.013	0.013
	Two Step (Flexible)	Mean	-0.216	0.108	0.108	-0.192	0.096	0.096	-0.127	0.063	0.064	-0.071	0.035	0.035	-0.038	0.019	0.019
		SD	0.044	0.022	0.022	0.045	0.023	0.023	0.027	0.015	0.015	0.017	0.010	0.011	0.016	0.011	0.012
		SE	0.063	0.025	0.025	0.048	0.021	0.021	0.025	0.012	0.012	0.016	0.010	0.010	0.019	0.015	0.015
	LCF	Mean	-0.204	0.102	0.102	-0.159	0.079	0.079	-0.108	0.054	0.054	-0.057	0.029	0.029	-0.012	0.006	0.006
		SD	0.027	0.015	0.014	0.023	0.012	0.012	0.020	0.011	0.010	0.021	0.012	0.011	0.026	0.014	0.014
		SE	0.025	0.013	0.013	0.021	0.011	0.011	0.019	0.010	0.010	0.020	0.011	0.011	0.024	0.013	0.013
	LCF (Flexible)	Mean	-0.169	0.084	0.085	-0.140	0.070	0.070	-0.108	0.054	0.054	-0.075	0.038	0.038	-0.046	0.023	0.023
		SD	0.035	0.018	0.018	0.026	0.014	0.013	0.019	0.011	0.010	0.019	0.011	0.011	0.025	0.014	0.014
		SE	0.032	0.016	0.016	0.024	0.013	0.013	0.019	0.010	0.010	0.017	0.010	0.010	0.021	0.012	0.012

1. APEs at $(\overline{z}_1, w_p)$.
2. $\pi_1 = 0$, $\rho = 1$.
3. The grey colored cells indicate that the estimates have the opposite directions to the true APEs.

Table 9: Average APEs under Condition 2 and χ^2_3 distribution

π_2			0.1			0.2			0.5			1		
g			1	2	3	1	2	3	1	2	3	1	2	3
w	True	Mean	-0.051	0.026	0.026	-0.051	0.025	0.025	-0.051	0.025	0.025	-0.049	0.025	0.025
	Two Step	Mean	-0.048	0.023	0.024	-0.045	0.022	0.023	-0.049	0.024	0.025	-0.047	0.024	0.024
		SD	0.093	0.056	0.053	0.046	0.026	0.026	0.015	0.009	0.008	0.007	0.004	0.004
		SE	-	-	-	-	-	-	-	-	-	-	-	-
	LCF	Mean	-0.019	0.016	0.003	-0.045	0.023	0.022	-0.053	0.027	0.027	-0.052	0.026	0.026
		SD	0.700	0.452	0.421	0.113	0.068	0.065	0.018	0.010	0.010	0.008	0.005	0.005
		SE	3.288	2.637	2.268	0.095	0.091	0.080	0.017	0.010	0.010	0.008	0.005	0.005
w^2	True	Mean	-0.062	0.031	0.031	-0.062	0.031	0.031	-0.062	0.031	0.031	-0.061	0.030	0.030
	Two Step	Mean	-0.060	0.029	0.031	-0.058	0.029	0.029	-0.061	0.030	0.030	-0.059	0.030	0.030
		SD	0.091	0.056	0.054	0.045	0.027	0.025	0.015	0.009	0.009	0.008	0.004	0.005
		SE	-	-	-	-	-	-	-	-	-	-	-	-
	Two Step	Mean	-0.045	0.022	0.024	-0.047	0.023	0.024	-0.057	0.029	0.029	-0.058	0.029	0.029
	(Flexible)	SD	0.094	0.056	0.056	0.052	0.029	0.029	0.015	0.009	0.009	0.008	0.004	0.004
		SE	-	-	-	-	-	-	-	-	-	-	-	-
	LCF	Mean	-0.044	0.016	0.028	-0.054	0.026	0.029	-0.065	0.033	0.033	-0.064	0.032	0.032
		SD	0.760	0.400	0.525	0.136	0.098	0.051	0.017	0.010	0.010	0.008	0.005	0.005
		SE	4.077	1.878	3.020	0.186	0.143	0.050	0.016	0.009	0.009	0.008	0.004	0.004
	LCF	Mean	-0.049	0.019	0.030	-0.051	0.024	0.027	-0.065	0.033	0.033	-0.064	0.032	0.032
	(Flexible)	SD	0.882	0.418	0.594	0.215	0.151	0.069	0.017	0.010	0.009	0.008	0.005	0.005
		SE	4.998	2.112	3.393	0.315	0.221	0.101	0.016	0.009	0.009	0.008	0.004	0.004

1. $\pi_1 = 0$, $\rho = 1$.

Table 10: Percentile APEs under Condition 2, $\pi_2 = 0.1$, and χ^2_3 distribution

w_p			10^{th}			25^{th}			50^{th}			75^{th}			90^{th}		
g			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
w	True	Mean	-0.097	0.049	0.049	-0.071	0.035	0.035	-0.043	0.022	0.022	-0.023	0.012	0.012	-0.012	0.006	0.006
	Two Step	Mean	-0.146	0.072	0.074	-0.098	0.048	0.050	-0.030	0.014	0.015	0.005	-0.002	-0.002	0.014	-0.007	-0.007
		SD	0.165	0.087	0.086	0.146	0.080	0.078	0.102	0.062	0.060	0.075	0.051	0.048	0.059	0.041	0.041
		SE	0.229	0.132	0.120	0.209	0.117	0.115	0.166	0.108	0.097	0.150	0.099	0.099	0.163	0.099	0.107
w^2	True	Mean	-0.146	0.073	0.073	-0.088	0.044	0.044	-0.043	0.022	0.022	-0.020	0.010	0.010	-0.010	0.005	0.005
	Two Step	Mean	-0.163	0.081	0.083	-0.125	0.061	0.064	-0.042	0.020	0.023	-0.002	0.000	0.002	0.008	-0.005	-0.003
		SD	0.174	0.092	0.090	0.158	0.086	0.084	0.100	0.064	0.061	0.069	0.050	0.049	0.052	0.040	0.040
		SE	0.220	17.102	13.753	0.209	7.810	7.516	0.196	2.186	3.149	0.170	0.184	0.250	0.208	1.094	5.276
	Two Step (Flexible)	Mean	-0.154	0.077	0.077	-0.105	0.052	0.054	-0.044	0.020	0.024	0.014	-0.008	-0.006	0.045	-0.023	-0.022
		SD	0.197	0.108	0.105	0.149	0.085	0.085	0.128	0.075	0.073	0.122	0.071	0.071	0.123	0.072	0.072
		SE	4E+29	7E+28	3E+25	2E+28	8E+27	2E+23	5E+22	3E+30	1E+27	2E+31	3E+29	1E+27	8E+27	5E+28	2E+30
		SE*	0.443	1.018	0.961	0.193	0.807	0.771	0.112	0.389	0.376	0.155	0.117	0.123	0.314	0.698	0.689
	LCF	Mean	-0.172	0.080	0.092	-0.111	0.050	0.062	-0.044	0.016	0.028	0.023	-0.017	-0.006	0.084	-0.047	-0.036
		SD	0.760	0.399	0.525	0.760	0.399	0.525	0.760	0.400	0.525	0.760	0.400	0.525	0.760	0.400	0.525
		SE	4.078	1.879	3.021	4.078	1.878	3.020	4.077	1.878	3.020	4.077	1.878	3.020	4.078	1.879	3.020
	LCF (Flexible)	Mean	-0.131	0.061	0.070	-0.093	0.041	0.051	-0.049	0.019	0.030	-0.006	-0.003	0.009	0.033	-0.023	-0.010
		SD	1.055	0.506	0.671	0.913	0.434	0.610	0.881	0.418	0.593	1.007	0.484	0.639	1.211	0.589	0.721
		SE	5.427	2.313	3.611	5.182	2.194	3.485	4.998	2.111	3.393	5.089	2.164	3.439	5.334	2.279	3.567

1. APEs at $(\overline{z}_1, w_p)$.

2. $\pi_1 = 0$, $\rho = 1$.

3. SE* is the median of the standard errors.

4. The grey colored cells indicate that the estimates have the opposite directions to the true APEs.

Table 11: Percentile APEs under Condition 2, $\pi_2 = 0.2$, and χ^2_3 distribution

w_p			10^{th}			25^{th}			50^{th}			75^{th}			90^{th}		
g			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
w	True	Mean	-0.098	0.049	0.049	-0.071	0.035	0.035	-0.043	0.022	0.022	-0.023	0.011	0.011	-0.012	0.006	0.006
	Two Step	Mean	-0.120	0.060	0.060	-0.070	0.035	0.035	-0.028	0.014	0.014	-0.009	0.004	0.005	-0.002	0.001	0.001
		SD	0.115	0.058	0.058	0.070	0.037	0.037	0.037	0.022	0.023	0.029	0.019	0.020	0.027	0.018	0.019
		SE	0.107	0.057	0.056	0.068	0.036	0.037	0.036	0.024	0.024	0.032	0.024	0.025	0.022	0.020	0.020
w^2	True	Mean	-0.146	0.073	0.073	-0.088	0.044	0.044	-0.043	0.022	0.022	-0.020	0.010	0.010	-0.010	0.005	0.005
	Two Step	Mean	-0.143	0.071	0.072	-0.094	0.047	0.048	-0.041	0.020	0.021	-0.016	0.008	0.008	-0.006	0.003	0.003
		SD	0.127	0.065	0.064	0.080	0.043	0.041	0.035	0.025	0.021	0.026	0.020	0.018	0.022	0.018	0.018
		SE	0.098	0.576	0.242	0.081	0.270	0.132	0.035	0.072	0.036	0.030	0.025	0.024	0.021	0.068	0.067
	Two Step (Flexible)	Mean	-0.139	0.069	0.070	-0.085	0.042	0.043	-0.038	0.019	0.020	-0.002	0.001	0.001	0.022	-0.011	-0.011
		SD	0.130	0.069	0.068	0.073	0.041	0.040	0.057	0.032	0.032	0.064	0.035	0.036	0.080	0.044	0.045
		SE	4E+06	2E+29	2E+27	1E+17	4E+29	1E+28	1E+29	4E+29	4E+28	6E+25	8E+18	2E+18	2E+25	5E+17	2E+17
		SE*	0.158	0.326	0.326	0.067	0.250	0.249	0.037	0.111	0.112	0.030	0.028	0.029	0.038	0.175	0.170
	LCF	Mean	-0.181	0.089	0.092	-0.121	0.059	0.062	-0.054	0.026	0.028	0.013	-0.008	-0.005	0.073	-0.038	-0.035
		SD	0.137	0.098	0.051	0.136	0.098	0.051	0.136	0.098	0.050	0.137	0.098	0.051	0.137	0.098	0.051
		SE	0.188	0.144	0.051	0.187	0.143	0.050	0.186	0.143	0.050	0.186	0.143	0.050	0.187	0.143	0.051
	LCF (Flexible)	Mean	-0.129	0.063	0.067	-0.092	0.044	0.048	-0.051	0.024	0.027	-0.010	0.003	0.007	0.027	-0.015	-0.012
		SD	0.284	0.186	0.106	0.242	0.166	0.083	0.215	0.151	0.069	0.212	0.146	0.074	0.231	0.151	0.091
		SE	0.381	0.254	0.133	0.347	0.236	0.116	0.315	0.221	0.101	0.307	0.220	0.100	0.333	0.234	0.113

1. APEs at $(\bar{z}_1, w_p)$.
2. $\pi_1 = 0$, $\rho = 1$.
3. SE* is the median of the standard errors.
4. The grey colored cells indicate that the estimates have the opposite directions to the true APEs.

Table 12: Percentile APEs under Condition 2, $\pi_2 = 0.5$, and χ^2_3 distribution

w_p			10^{th}			25^{th}			50^{th}			75^{th}			90^{th}		
g			1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
w	True	Mean	-0.098	0.049	0.049	-0.070	0.035	0.035	-0.042	0.021	0.021	-0.022	0.011	0.011	-0.011	0.006	0.006
	Two Step	Mean	-0.100	0.050	0.050	-0.065	0.032	0.032	-0.036	0.018	0.018	-0.019	0.010	0.010	-0.011	0.005	0.006
		SD	0.045	0.023	0.023	0.023	0.013	0.012	0.008	0.006	0.006	0.003	0.005	0.005	0.003	0.005	0.005
		SE	0.042	0.021	0.021	0.021	0.012	0.011	0.007	0.006	0.006	0.003	0.005	0.005	0.003	0.005	0.005
w^2	True	Mean	-0.147	0.074	0.074	-0.087	0.044	0.044	-0.042	0.021	0.021	-0.019	0.010	0.010	-0.009	0.005	0.005
	Two Step	Mean	-0.128	0.064	0.064	-0.088	0.044	0.044	-0.047	0.024	0.024	-0.022	0.011	0.011	-0.010	0.005	0.005
		SD	0.058	0.029	0.029	0.028	0.015	0.015	0.010	0.007	0.007	0.006	0.006	0.006	0.004	0.006	0.006
		SE	0.047	0.059	0.059	0.024	0.032	0.032	0.009	0.010	0.010	0.005	0.006	0.006	0.005	0.028	0.028
	Two Step (Flexible)	Mean	-0.129	0.064	0.065	-0.081	0.041	0.041	-0.042	0.021	0.021	-0.019	0.010	0.009	-0.008	0.004	0.004
		SD	0.056	0.029	0.029	0.025	0.014	0.014	0.013	0.008	0.008	0.009	0.007	0.007	0.010	0.009	0.009
		SE	0.058	0.093	0.093	0.029	0.063	0.063	0.025	0.031	0.031	0.043	0.018	0.019	0.110	0.068	0.069
	LCF	Mean	-0.186	0.093	0.093	-0.129	0.064	0.064	-0.065	0.033	0.033	-0.002	0.001	0.001	0.055	-0.028	-0.028
		SD	0.025	0.014	0.013	0.020	0.011	0.011	0.017	0.010	0.010	0.017	0.010	0.010	0.020	0.011	0.011
		SE	0.022	0.012	0.012	0.019	0.010	0.010	0.016	0.009	0.009	0.016	0.009	0.009	0.019	0.011	0.011
	LCF (Flexible)	Mean	-0.141	0.070	0.071	-0.105	0.052	0.053	-0.065	0.033	0.033	-0.025	0.013	0.012	0.011	-0.005	-0.006
		SD	0.045	0.023	0.023	0.031	0.016	0.016	0.017	0.010	0.009	0.012	0.008	0.008	0.021	0.013	0.013
		SE	0.039	0.021	0.021	0.028	0.015	0.015	0.016	0.009	0.009	0.010	0.007	0.007	0.016	0.011	0.011

1. APEs at $(\overline{z}_1, w_p)$.

2. $\pi_1 = 0$, $\rho = 1$.

3. The grey colored cells indicate that the estimates have the opposite directions to the true APEs.

Table 13: Mean Squared Errors of Average APEs estimates under Condition 2 and Normal distribution

π_2		0.1			0.2			0.5			1		
g		1	2	3	1	2	3	1	2	3	1	2	3
w	two step	0.011	0.004	0.003	0.004	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000
	LCF	1.214	0.535	0.183	0.154	0.036	0.042	0.001	0.000	0.000	0.000	0.000	0.000
w^2	two step	0.012	0.004	0.004	0.004	0.001	0.001	0.001	0.000	0.000	0.000	0.000	0.000
	two step (Flexible)	0.012	0.004	0.004	0.004	0.001	0.001	0.001	0.000	0.000	0.000	0.000	0.000
	LCF	1.592	0.341	0.496	0.160	0.052	0.030	0.001	0.000	0.000	0.000	0.000	0.000
	LCF (Flexible)	1.908	0.423	0.571	0.220	0.073	0.040	0.001	0.000	0.000	0.000	0.000	0.000

1. The Mean Squared Errors are calculated from Table 5.

2. $\pi_1 = 0$, $\rho = 1$.

Table 14: Mean Squared Errors of Percentile APEs estimates under Condition 2, $\pi_2 = 0.1$, and Normal distribution

w_p		10^{th}			25^{th}			50^{th}			75^{th}			90^{th}		
g		1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
w^2	two step	0.018	0.005	0.005	0.032	0.009	0.009	0.019	0.006	0.006	0.008	0.003	0.003	0.005	0.002	0.002
	two step (Flexible)	0.024	0.008	0.007	0.032	0.009	0.009	0.021	0.006	0.007	0.013	0.004	0.005	0.012	0.004	0.004
	LCF	1.594	0.342	0.497	1.594	0.342	0.497	1.593	0.341	0.497	1.592	0.341	0.496	1.591	0.341	0.496
	LCF (Flexible)	2.184	0.480	0.654	1.981	0.436	0.595	1.900	0.422	0.568	2.022	0.456	0.595	2.270	0.521	0.655

1. The Mean Squared Errors are calculated from Table 6.

2. $\pi_1 = 0$, $\rho = 1$.

Table 15: Four levels of MEAP

Variable	Description
Level 1	Exceeded Michigan Standards
Level 2	Met Michigan Standards
Level 3	demonstrated Basic knowledge and skills of Michigan Standards
Level 4	Apprentice level, showing little success in meeting Michigan standards

Table 16: The descriptions for the dependent variables

Variable	Mean	SD	Min	Max	Description
y_1	0.282	0.139	0	1.000	fraction of Level 1
y_2	0.463	0.084	0	0.742	fraction of Level 2
y_3	0.222	0.099	0	0.643	fraction of Level 3
y_4	0.033	0.038	0	0.329	fraction of Level 4
Total	1.000				

Table 17: Sample mean of the right-hand side variables

Variable	Mean (standard deviations)
enrollment	3132.272 (6939.982)
Fraction of applications for free and reduced lunch	0.353 (0.178)
per pupil expenditure	8092.164 (1092.165)
foundation allowance	6982.72 (655.638)
# of districts	518

Table 18: The first step estimation

	coefficient	SE	t	p-values
log(enrollment)	0.006	0.005	1.190	0.233
lunch	0.306	0.025	12.470	0.000
log(foundation allowance)	1.069	0.061	17.510	0.000
constant	-0.617	0.528	-1.170	0.244
R^2	0.6111			
$F(H_0 :$ log(foundation allowance)=0)	306.63			

Table 19: Average APEs estimates of log(spending) on 4^{th} grade MEAP math test

level			1	2	3	4
w	Two Step	Estimates	0.242*	-0.166*	-0.069	-0.007
		SE	0.050	0.033	0.037	0.017
	LCF	Estimates	0.275*	-0.207*	-0.069*	0.002
		SE	0.057	0.035	0.034	0.014
w^2	Two Step	Estimates	0.287*	-0.197*	-0.096	0.006
		SE	0.070	0.049	0.055	0.018
	Two Step (Flexible)	Estimates	0.266*	-0.189*	-0.067	-0.010
		SE	0.076	0.050	0.057	0.021
	LCF	Estimates	0.319	-0.220	-0.101	0.003
		SE	2135.4	559.4	1282.7	293.2
	LCF (Flexible)	Estimates	0.297*	-0.207*	-0.071	-0.019
		SE	0.092	0.053	0.057	0.021

1. w = log(spending).
2. The standard errors of the pseudo/pseudo flexible APE estimates are calculated using 1000 bootstrap replications.
3. The standard errors of the LCF estimates are adjusted.
4. * is significant at, or below, 5 percent.

Table 20: Percentile APEs estimates of log(spending) on 4^{th} grade MEAP math test

w_p			10^{th}				25^{th}				50^{th}			
level			1	2	3	4	1	2	3	4	1	2	3	4
w	Two Step	Estimates	0.229*	-0.158*	-0.065	-0.006	0.234*	-0.161*	-0.067	-0.006	0.241*	-0.165*	-0.069	-0.007
		SE	0.053	0.038	0.042	0.017	0.055	0.038	0.042	0.017	0.059	0.038	0.042	0.017
w^2	Two Step	Estimates	0.317	-0.215	-0.123	0.022	0.311	-0.212	-0.117	0.017	0.300*	-0.205	-0.106	0.011
		SE	0.226	0.310	0.185	0.038	0.174	0.222	0.132	0.034	0.115	0.170	0.082	0.038
	Two Step	Estimates	0.251*	-0.202*	-0.047	-0.002	0.257*	-0.199*	-0.054	-0.004	0.265*	-0.193*	-0.064	-0.007
	(Flexible)	SE	0.089	0.079	0.079	0.035	0.087	0.073	0.072	0.056	0.083	0.062	0.062	0.080
	LCF	Estimates	0.368	-0.235	-0.137	0.003	0.354	-0.231	-0.126	0.003	0.331	-0.224	-0.110	0.003
		SE	4481.9	1174.2	2692.3	615.4	3789.0	992.7	2276.1	520.2	2694.5	705.9	1618.6	370.0
	LCF	Estimates	0.295	-0.202	-0.061	-0.032	0.296*	-0.203*	-0.064	-0.028	0.296*	-0.206*	-0.069	-0.022
	(Flexible)	SE	0.152	0.089	0.095	0.040	0.133	0.077	0.083	0.034	0.105	0.061	0.065	0.025

w_p			75^{th}				90^{th}			
level			1	2	3	4	1	2	3	4
w	Two Step	Estimates	0.251*	-0.171*	-0.073	-0.007	0.265*	-0.180*	-0.077	-0.008
		SE	0.064	0.038	0.042	0.016	0.071	0.040	0.041	0.016
w^2	Two Step	Estimates	0.279*	-0.190	-0.089	0.001	0.238	-0.162	-0.063	-0.014
		SE	0.126	0.207	0.086	0.054	0.230	0.267	0.117	0.081
	Two Step	Estimates	0.276*	-0.186*	-0.078	-0.012	0.291*	-0.176*	-0.097	-0.018
	(Flexible)	SE	0.075	0.049	0.056	0.085	0.077	0.044	0.069	0.055
	LCF	Estimates	0.299	-0.214	-0.087	0.002	0.250	-0.200	-0.051	0.002
		SE	1156.1	302.9	694.5	158.7	1184.8	310.4	711.7	162.7
	LCF	Estimates	0.297*	-0.209*	-0.075	-0.013	0.298*	-0.214*	-0.085*	0.001
	(Flexible)	SE	0.072	0.042	0.044	0.015	0.067	0.045	0.039	0.021

1. w = log(spending).
2. The standard errors of the pseudo/pseudo flexible APE estimates are calculated using 1000 bootstrap replications.
3. The standard errors of the linear CF estimates are adjusted.
4. * is significant at, or below, 5 percent.