Nonparametric Slope Estimators for Fixed-Effect Panel Data

Kusum Mundra[1]
Department of Economics
San Diego State University
San Diego, CA 92182

October 2004

[1] I would like to acknowledge helpful comments from Aman Ullah. This version also includes many constructive suggestions received from participants at the 11th International Panel Data Conference, in July 2004 at the Texas A&M University.

Abstract

There are more data sets being available in panel form, msking panel data estimation techniques very popular in applied econometrics. Dependent on how the sample is drawn from the population, the cross sectional effect is treated as fixed or random. In this paper, I present two nonparanetric slope estimation for fixed effect panel models. It is will known in theoretical econometrics that misspecification of the functional form leads to biased estimates of the parameters. In particular, there has been extensive work on semiparametric or more appropriately partially linear models, however the literature does not talk enough on nonparametric panel models. Recently there has been some work in the nonparametric estimation of random effect models, see Henderson and Ullah (2003). Unfortunately the literature does not talk about nonparametric fixed-effect slope estimation, in this paper I intent to fill that gap. Two nonparametric slope estimates for fixed effect panel models are discussed with the asymptotics established. Finite sample Monte Carlo properties are discussed. I apply the nonparametric estimators to investigate the effect of age and tenure on the workers hourly wages using NLSY79 (National Longitudinal Survey of Youth Data). In estimating earning functions it is a very common practice to assume workers earnings being quadratic in age and tenure, the estimation shows that the parametric slope shows a steady decline in hourly earnings as age goes up but the nonparametric slope shows that the decline is not steady. In addition, the nonparametric estimator shows that for bigger magnitude of tenure the effect on the hourly wages is quite different from the lower weeks of tenure.

1 Introduction

More data sets are being available in panel form, making panel data estimation techniques very popular in applied econometrics. In panel data we are able to get observations over cross-sectional units over time and we can capture the cross-ssectional heterogeneity by including an individual or cross-sectional effect in our model. Dependent on how the sample is drawn, the cross sectional effect is treated as fixed or random. In the random effect specification it is assumed that the cross section is drawn from a random population and the cross sectional effect is part of the stochastic error. Whereas in the fixed effect model, the cross sectional effect capturing the cross-section heterogeneity is not a part of the error but it is a parameter varying across the cross-section. The cross sectional effect when treated as fixed and non-random, allows the cross sectional effect for the cross section i to be correlated with other exogenous regressors in the model. [1]In this paper, we present two nonparanetric slope estimation for fixed effect panel models.

It is will known in theoretical econometrics that misspecification of the functional form leads to biased estimates of the parameters. and often policies are based on these biased estimates, making this a significant problem in applied econometrics. The significance of functional form in the modelling makes nonparametric analysis very important. In nonparametric models, no specific functional form is imposed on how does the indepenedent variable affects the dependent variable, see Ullah and Pagan (1999). An important extension of nonparametric Kernel techniques have been to panel data models, see Ullah and Roy (1998), Porter (1996). In particular, there has been extensive work on semiparametric or more appropriately partially linear models (where some dependent variable in the regression model enter linearly and for others the functional form is not known, and hence partially linear), see Li and Stengos (1996), Li and Ullah (1998) and Berg, Li and Ullah (2000). However, there has not been enough work on nonparametric panel models, recently there has been some work in the nonparametric estimation of random effect models, see Henderson (2004). Unfortunately the literature does not talk about nonparametric fixed-effect slope estimation, in this paper we intent to fill that gap. This paper presents two

[1] For discussion on different panel data estimation and treatment of the cross-sectional effect see Hsiao (2003), Baltagi (2002) and Mundlak (1978).

nonparametric slope estimates for fixed effect panel model. The estimators are applied to estimate earning functions using NLSY data.

This paper is organized as follows. Section 2 specifies the model and gives the estimates and section 3 establishes the asymptotics of the estimates. Monte Carlo results are discussed in section 4. Section 5 presents the application of the two estimates to earning function using NLSY data. Finally, section 6 concludes.

2 Nonparametric Slope Estimation

The parametric (linear) fixed-effcet panel model is specified as follows:

$$y_{it} = \alpha_i + z_{it}\beta + u_{it} \quad i = 1, ..., n \ t = 1, ..., T \tag{2.1}$$

where y_{it} is the dependent variable, z_{it} is a matrix of k exogenous variable and β is a $k \times 1$ vector and α_i is the cross-sectional effect that is treated as non-random and is a fixed unknown parameter to be estimated. The error u_{it} is assumed to follows the usual *iid* error structure with mean zero and constant variance. The nonparametric model given in (2.1) with the fixed effect is as follows

$$y_{it} = \alpha_i + m(z_{it}) + u_{it} \quad i = 1, ..., n \ t = 1, ..., T \tag{2.2}$$

where we do not specify how z_{it} effects $y_{it,}$ the unknown functional form $m(.)$ makes the model a nonparametric model. The problem is to estimate β (the parametric slope) in the model (2.1) nonparametrically in (2.2). The nonparametric approach is to use the nonparametric kernel regression estimation of the unknown form $m(z_{it})$ and estimate $m'(z_{it})$, where $m'(z_{it})$ is the first derivative of $m(z_{it})$ with respect to z_{it}. The model in (2.2) can be written as

$$y_{it} = \alpha_i + m(z) + (z_{it} - z)\beta(z) + (1/2)(z_{it} - z)^2 m^2(z) + u_{it} \tag{2.3}$$

where we expand the unknown regression around a point z, to the third order, the idea being in (2.3) to estimate the slope $m'(z_{it})$ in (2.2) locally in the interval h around z by linear approximation $(z_{it} - z)\beta(z)$.[2]

[2]This is similar to the nonparametric kernel regression functional, varying coefficient models proposed by Lee and Ullah (2003), Cai et al. (2000).

There are two well known transformations used to take care of the fixed effect in the parametric models, one is first differencing $y_{it} - y_{it-1}$ and the second is taking deviations from mean $y_{it} - \bar{y}_{i.} = y_{it} - \frac{1}{T}\sum y_{it}$, see Hsiao (2003), Baltagi (2002), Chamberlain (1984), Matyas and Sevestre (1996). In the linear parametric panel model little is known how these two methods compare against each other, in section 4 of this paper I present some Monte Carlo simulations results for the two transformations. In this paper, we use the two transformations to the fixed effect nonparametric panel data models and estimate the slope coefficients with local linear kernel weighted techniques. In the first estimator, we use the first differencing transformation to take account of the cross-sectional effect while estimatig the slope parameter of interest. In the second estimator deviation from mean is used.

2.1 First - Differencing Estimator

After taking a first difference of (2.3) we get:

$$\Delta y_{it} = (z_{it} - z_{it-1})\beta(z) + \frac{1}{2}[(z_{it} - z)^2 - (z_{it-1} - z)^2]m^2(z) + \Delta u_{it} + r \tag{2.4}$$

where $\beta(z) = m^1(z)$ is the slope parameter of interest and r is the remainder term. The model is (2.4) can also be written as

$$A_2 Y \tilde{} A_2 Z\beta(z) + A_2 U$$

where A_2 is $\begin{pmatrix} -1 & 1 & 0 & . & . & . & 0 \\ 0 & -1 & 1 & . & . & . & 0 \\ . & . & . & . & . & . & . \\ 0 & 0 & . & . & . & -1 & 1 \end{pmatrix}$ matrix of order n(T-1) x nT.

The local linear estimator of is the parameter of interest and is given by,

$$\hat{\beta}(z) = (Z'A_2KA_2Z)^{-1}Z'A_2KA_2Y \tag{2.5}$$

or

$$\hat{\beta}(z) = \sum_{i=1}^{n}\sum_{t=2}^{T} w_{it}\Delta y_{it} \tag{2.6}$$

where $w_{it} = \frac{\Delta z_{it} K_{it} K_{it-1}}{\Sigma\Sigma\Delta^2 z_{it} K_{it} K_{it-1}}$, see Pagan and Ullah (1999). Where $K_{it} = K(\frac{z_{it}-z}{h})$ and $K_{it-1} = K(\frac{z_{it-1}-z}{h})$ are the standard normal kernel function with optimal window width h.[3]

.

2.2 Deviation from Mean Estimator

Deviation from mean transformation for the panel data model is proposed as follows[4]:

$$\bar{y}_{i.} = \alpha_i + m(z) + (\bar{z}_{i.} - z)\beta(z) + \overline{u}_{i.} + r \tag{2.7}$$

where $\bar{y}_{i.} = \frac{1}{T}\sum y_{it}$, $\bar{z}_{i.} = \frac{1}{T}\sum z_{it}$, and $\overline{u}_{i.} = \frac{1}{T}\sum u_{it}$. Taking a difference of (2.7) from (2.1) gives

$$y_{it} - \bar{y}_{i.} = (z_{it} - \bar{z}_{i.})\beta(z) + u_{it} - \overline{u}_{i.} + r$$

The local FE estimator of the slope $\beta(z)$ can then be obtained by minimizing $\underset{i}{\Sigma}\underset{t}{\Sigma}(y_{it} - \bar{y}_{i.} - (z_{it} - \bar{z}_{i.})\beta(z))^2 K(\frac{z_{it}-z}{h})$.This gives, for q = 1,the slope estimator as follows:

$$\tilde{\beta}(z) = \sum_i \sum_t \frac{K_{it}(y_{it} - \bar{y}_{i.})(z_{it} - \bar{z}_{i.})}{\underset{i}{\Sigma}\underset{t}{\Sigma} K_{it}(z_{it} - \bar{z}_{i.})^2}, \tag{2.8}$$

and for q $\geq$ 1,

$$\tilde{\beta}(z) = (X' M_D K(z) M_D X)^{-1} X' M_D K(z) M_D y$$

where $M_D = I_{NT} - D(D'D)^{-1}D' = I_{NT} \otimes Q_T$ and $Q_T = I_T - \bar{Q}_T$, $\bar{Q}_T = \frac{\iota_T \iota_T'}{T}$

3 Asymptotic Properties of the Estimators

In this section asymptotic properties of the estimators are established and asymptotic distributions of the above estimators are derived. The assumptions and steps are similar to those of Robinson (1986, 1988a, 1988b), Kneisner and Li (2002)

[3]See Pagan and Ullah (1999) for well established properties of the standard normal kernel and details on the optimal window width (bandwidth) selection.

[4]In Ullah and Roy (1997) the mean deviation nonparametric fixed effect estimator was mentioned but the properties of the estimator were not discussed.

Following Robinson (1988) let G_{μ}^{λ} denote the class of functions such that if $g \epsilon G_{\mu}^{\lambda}$, then g is μ times differentiable; g and its derivatives (up to order μ) are all bounded by some function that has $\lambda - th$ order finite moments. Also, K_2 denotes the class of non-negative kernel functions k : satisfying $\int k(v)\, v^m\, dv = \delta om$ for $m = 0\,,\, 1$(δom is the Kronecker's delta), $\int k(v) v v' dv = C_k I(I > 0\,,$ and $k(u) = O((1+ \mid u \mid^{3+\eta})^{-1})$ for some $\eta > 0$. Further, $\int k^2(v) v v'\, dv = D_k\, .I$.

Property 1: Under the following assumptions

(1) (i) for all t, $(y_{it}\,,\, z_{it})$ are i.i.d. across $i.z_{it}$ is a strictly stationary real valued stochastic process and z_{it} and z_{it-1} admits a joint density function $f \epsilon G_{\mu-1}^{\infty}$. $m(z_{it})$ and $m(z_{it-1})$ both $\epsilon\, G_{\mu-1}^{4}$ for some positive integer $\mu > 2$.

(2) $E(u_{it} \mid z_{it},\, z_{it-1}) = 0\,,\, E(u_{it}^2 \mid z_{it}, z_{it-1}) = \sigma^2$ is continuous in z_{it} and z_{it}, and u_{it},

(3) $k \,\epsilon\, K_2$ and $k(v) \geq 0\,;$ as $n \to \infty\,,\, h \to 0\,,\, nh^{q+3} \to \infty$ and $nh^{q+4} \to 0$.

$$\sqrt{nTh^{q+3}}\left(\hat{\beta}(z) - \beta(z)\right) \tilde{\ } N\left(0, \Sigma\right)$$

for large N and fixed T, where $R \simeq m^2(z)\,(\mu_2 f(z,z))\, h{-}1$, $\Phi = 4\sigma^2 \mu_2 f\,(z,z)\,(\mu_2 - 1)\,,$ $\Sigma = R^{-1} \Phi R^{-1}$

For the derivation of the property 1 see Appendix A.

Property 2: Under the following assumptions

(1) (i) for all t, $(y_{it}\,,\, z_{it})$ are i.i.d. across i and z_{it} admits a density function $f \epsilon G_{\mu-1}^{\infty}$, $m(z_{it})\, \epsilon\, G_{\mu-1}^{4}$ for some positive integer $\mu > 2$.

(2) $E(u_{it}|z_{it}) = 0\,,\, E(u_{it}^2 \mid z_{it}) = \sigma_u^2$ is continuous in z_{it}, and u_{it},

(3) $K \,\epsilon\, K_2$ and $k(v) \geq 0\,;$ as $n \to \infty\,,\, h \to 0\,,\, nh^{q+2} \to \infty$ and $nh^{q+3} \to 0$.

$$\sqrt{nTh^{q+2}}\left(\tilde{\beta}(z) - \beta(z)\right) \tilde{\ } N\left(0, \Sigma_1\right)$$

where $\Sigma_1 = R_1^{-1} \Phi_1 R^{-1}$, where $R_1 = T^2 z^2 f(z)$ and $\Phi_1 = \frac{(T-1)}{T} \sigma_u^2 [\int K^2\,(\psi_{1t})\, \psi_{1t}^2 d\psi_{1t}]$.

For the derivation of the property 2 see, Appendix B.

4 Monte Carlo Results

In this section we discuss the Monte Carlo properties of the above estimators for different sizes of sample, to investigate the finite sample properties.[5] Firstly, the monte carlo properties is investigated and compared between the two prametric linear fixed effect models. Secondly in this section, we present the above nonparametric fixed effect estimators.

For the parametric linear model the following data generating process is used is

$$y_{it} = \alpha_i + z_{it}\beta + u_{it} \tag{4.1}$$

where α_i is the cross sectional fixed effect and is generated by $\alpha_i = 2.5 + \alpha_j$, this allows that the fixed effect for unit i is correlated with j. The above model (where the data generation is linear) is estimated by both the methods, deviation from mean and first differencing for different N and T. The $Bias(\widehat{\widehat{\beta}}) = M^{-1}\sum_j \left(\widehat{\widehat{\beta}}_j - \beta\right)$ and $RMSE(\widehat{\widehat{\beta}}) = \left\{M^{-1}\sum_j \left(\widehat{\widehat{\beta}}_j - \beta\right)^2\right\}^{-1/2}$, where M is the number of replications, β_j is the estimate of β at the $j-th$ replication using NT observations. $M = 2000$ in the parametric simulation and T is varied to be 3, 6, 10, 50,100, and 500. While N takes the values 10, 50, 100, 500, and 1000. The parametric estimators are as follows:

(1) Parametric first differencing estimator

$$\widehat{\beta}_{diff} = \frac{\sum\sum\left(z_{it} - z_{it-1}\right)\left(y_{it} - y_{it-1}\right)}{\sum\sum(z_{it} - z_{it-1})^2}$$

(2) Parametric mena deviation estimator

$$\widehat{\beta}_{dev} = \frac{\sum\sum\left(z_{it} - \overline{z}_{i.}\right)\left(y_{it} - \overline{y}_{i.}\right)}{\sum\sum\left(z_{it} - \overline{z}_{i.}\right)^2}$$

Following the methodology of Baltagi, Chang and Li (1992), Li and Ullah (1992) x_{it} is generated by the process used by Nerlove (1971), where $z_{it} = o.1t + 0.5z_{it-1} + w_{it}$, where $z_{io} = 10 + 5w_{io}$ and $w_{it} \sim U[-0.5, 0.5]$, u_{it} is drawn from standard normal distribution and β is chosen to be 8. The simulation results showing the difference between the RMSE of the two parametric slope estimators are given in Table 1 (Panel

[5] This Section is work in progress.

A). In another exercise the data is generated by the method followed in Berg, Li and Ullah (.) where $z_{it} \sim U[-\sqrt{3}, \sqrt{3}]$, the results are given in Table 1 (Panel B). From Table 1 (both Panel A and Panel B) we see that for all N as T increases first differencing fixed effect slope estimator for linear model is doing better than the mean-deviation estimator. We see that the difference between the root mean square of the first-differencing and the mean deviation estimator is steadily rising as T goes up. The bias and the standard error for the first differencing etsimator is lower than the mean-deviation etsimator as T increases for all N.

For the nonparametric model the following data generating process is used is

$$y_{it} = \alpha_i + z_{it}\beta_1 + z_{it}^2\beta_2 + u_{it} \tag{4.2}$$

where $z_{it} \sim U[-0.5, 0.5]$ by Berg. Li and Ullah method and α_i is generated by $\alpha_i = v_i + c_2\alpha_j$, where $v_i \tilde{} N(0, \sigma_v)$, $M = 1000$ for the nonparametric simulations. The value of β_1 is chosen to be 0.5, β_2 is chosen to be 2, and c_1 is chosen to be 2. The value of $\sigma_v^2 + \sigma_u^2 = 20$ and $\rho = \sigma_u^2/(\sigma_v^2 + \sigma_u^2)$ takes the value of 0.8. In the above model the true data generation is quadratic and the model is estimated by both the nonparametric methods proposed in the previous sections; deviation from mean and first differencing. T is varied to be 3,6,10, while N takes the values 10, 50, 100. For comparison purposes we also compute the parametric fixed effect slope estimator for the model given in (4.2) by the differencing ($\widehat{\beta}_{diff}$) and the mean deviation ($\widehat{\beta}_{dev}$) estimator. Both in the case of the differencing transformation (Table 2) and in the case of the mean deviation transformation (Table 3) compared to the parametric estimators nonparametric estimator is doing better $\forall$ N and T. Also, we see that for fixed T and increasing N in both the estimators the difference in the $RMSE$ is falling between the parametric and the nonparametric etsimators.

Treating the cross sectional effect as the fixed effect allows that the individual effect maybe correlated with one or more of the explanatory variables (is very common when the model is misspecified, Hsiao (2003). In order to investigate the difference between the RMSE for the nonparametric and the parametric etsimator we carry out the above monte carlo simulation α_i to be correlated with $\overline{z}_{i.}$ by $\alpha_i = v_i + c_1\overline{z}_{i.} + c_2\alpha_j$, where the value of $c_1 = 2$ and $c_2 = 2$, results from the simulation are given in Table 4 and Table 5. Compared to the cases where α_i is not correlated with $\overline{z}_{i.}$, we see that the difference between the $RMSE$ of the parametric and the nonparametric estimator has gone up when we allow the fixed effect to be correlated with the explanatory variable.

We also incraese the degree of nonlinearity in the model given by (4.2), by increasing the value of β_2 from 2 to 4 and $\alpha_i = v_i + c_2\alpha_j$, where c_2= 2. From Table 6 and Table 7 we see that in small samples the nonparametric estimator is doing better.

5 Application

In this part, we apply the nonparametric mean deviation estimator to investigate the effect of age and tenure on the workers hourly wages using NLSY79 (National Longitudinal Survey of Youth Data). This a well known panel data that uses surveys by the Bureau of Labor Statistics (BLS) to gather information on the labor market experiences of diverse groups of men and women in the U.S. at different tim points.[6] In estimating earning functions it is a very common practice to assume workers earnings being quadratic in age and tenure, see Angrist and Krueger (1991), Sander (1992), Vella and Verbeek (1998) and Rivera-Batiz (1999) to name a few.

Using the fixed effect mean-deviation estimator the earnings slope is estimated with respect to age and tenure uisng a sample of 1000 individuals for $t = 3$ (the years are 1994, 1996, and 1998). For the parametric model the earning function is assumed to be quadratic in age (measured in years) and tenure (measured in weeks). The slope estimates are presented for age in Figure 1 and for tenure in Figure 2, respectively. Figure 1 shows that for all the individuals in the sample the parametric slope (shows a steady decline in hourly earnings as age goes up but the nonparametric slope shows that the decline is not steady. Also, the magnitude of the parametric results is very high compared to the nonparametric results. The parametric effect of tenure on wage shows that an additional gain of a week has close to zero effect on the hourly earnings. The nonparametric estimator shows that for bigger magnitude of tenure the effect on the hourly wages is quite different from the lower weeks of tenure.

[6]The NLS contractors for the BLS are the Centre for Human Resource Research (CHRR) at the Ohio State University, The National Opinion Research Center at the University of Chicago, and the U.S. Census Bureau.

Figure 3 and Figure 4 shows the nonparametric slope for age and tenure respectively through first-differencing transformation.

6 Appendix

6.1 Proof of Theorem 1

For $q = 1$ as $\hat{\beta}(z) = (Z'A_2KA_2Z)^{-1}Z'A_2KA_2Y = \sum_{i=1}^{n}\sum_{t=2}^{T} w_{it}\Delta y_{it}$

where $w_{it} = \frac{\Delta z_{it}K_{it}K_{it-1}}{\Sigma\Sigma\Delta^2 z_{it}K_{it}K_{it-1}}$. Refer to (2.5). This proof is for $q = 1$, but can easily be generalized.

Write, $E(\hat{\beta}(z)/z_{it},\ z_{it-1})\ = E(\sum\sum w_{it}\,(m(z_{it}) - m(z_{it-1})))$

. Approximated value is $E(\hat{\beta}(z)\ /\ z_{it},\ z_{it-1})$

$$\tilde{}\ E\left(\sum\sum w_{it}\left(\begin{array}{c}\Delta z_{it}\beta(z) + \frac{1}{2}\left[(z_{it}-z)^2 - (z_{it-1}-z)^2\right]m^2(z)\\ +\frac{1}{6}\left[(z_{it}-z)^3 - (z_{it-1}-z)^3\right]m^3(z)\end{array}\right)\right).$$

Using $\sum\sum \Delta z_{it}w_{it} = 1$

the approximated bias is, $E(\hat{\beta}(z)/z_{it},\ z_{it-1}) - \beta(z) =$

$E\left(\frac{1}{2}w_{it}m^2(z)\left[(z_{it}-z)^2 - (z_{it-1}-z)^2\right] + \frac{1}{6}m^3(z)w_{it}\left[(z_{it}-z)^3 - (z_{it-1}-z)^3\right]\right)$

Using $\psi_{it} = \frac{z_{it}-z}{h}$, we derive some lemmas.

$$\begin{aligned} E(D^1) &= E[\sum_i\sum_t\Delta^2 z_{it}K_{it}K_{it-1}] == nT\iint \Delta^2 z_{it}K_{it}K_{it-1}f(z_{it}, z_{it-1})dz_{it}dz_{it-1} \\ &\quad \tilde{}\, nTh^4 2\mu_2 f(z,z) + O(h^6) \end{aligned}$$

$$\begin{aligned} E(D^2) &= E[\Sigma\Sigma(z_{it}-z)^2\Delta z_{it}K_{it}K_{it-1}] \\ &= = nT\int\int h^4(\psi_{it})^2 K(\psi_{it})K(\psi_{it-1})f(\psi_{it}h+z, \psi_{it-1}h+z)d\psi_{it}d\psi_{it-1} \\ &\quad \tilde{}\, nT\left[-h^5\mu_2 f(z,z) + h^6\mu_4 f_{10}(z,z) - h^6(\mu_2)^2 f_{01}(z,z)\right] + O(h^7) \end{aligned} \tag{A.2}$$

$$\begin{aligned} E(D^3) &= E[\Sigma\Sigma(z_{it-1}-z)^2\Delta z_{it}K_{it}K_{it-1}] \\ &= = nT\int\int h^4(\psi_{it-1})^2 K(\psi_{it})K(\psi_{it-1})f(\psi_{it}h+z, \psi_{it-1}h+z)d\psi_{it}d\psi_{it-1} \\ &\quad \tilde{}\, nT\left[h^5 f(z,z)\mu_2 - h^6\mu_4 f_{01}(z,z) + 2h^6(\mu_2)^2 f_{10}(z,z)\right] + O(h^7) \end{aligned} \tag{A.3}$$

$$E(D^4) = E[\underset{i}{\Sigma}\underset{t}{\Sigma}\Delta z_{it}\Delta u_{it}K_{it}K_{it-1}] == 0$$

where the notation $f_{10}[x,y]$ represents the partial derivative of $f(x,y)$ with respect to the first variable

and $f_{01}[x,y]$ represents the partial derivative of $f(x,y)$ with respect to the second variable, and $f(z,z)$

is the value of $f(x,y)$ evaluated at $x = z$, $y = z$.

Combining (A.1-A.3) the approximate bias is: $E(\hat{\beta}(z)/z_{it},\ z_{it-1}) - \beta(z) = -\frac{1}{2}m^2(z)h + O(h^6)$.

Since approximate bias is free from z_{it}, z_{it-1} it is also approximate unconditional bias.

$$\begin{aligned}
\hat{\beta}(z) - \beta(z) &= \frac{m^2(z)}{2}\left[\frac{\sum\sum \Delta z_{it}K_{it}K_{it-1}\left[(z_{it}-z)^2 - (z_{it-1}-z)^2\right] + \Delta u_{it}}{\Sigma\Sigma\Delta^2 z_{it}K_{it}K_{it-1}}\right] \\
&= \frac{m^2(z)}{2}\left[D^1\right]^{-}\{\sum\sum \Delta z_{it}K_{it}K_{it-1}(z_{it}-z)^2 - \sum\sum \Delta z_{it}K_{it}K_{it-1}(z_{it-1}-z)^2 \\
&\quad \sum\sum \Delta z_{it}K_{it}K_{it-1}\Delta u_{it}\} \\
&= \frac{m^2(z)}{2}\left[D^1\right]^{-1}\left[D^2 + D^3 + D^4\right]
\end{aligned}$$

Now using lemmas A.1-A.4:

$E\left(\frac{D^1}{NTh^4}\right) = 2\mu_2 f(z,z) + o(1)$, thus $\frac{m^2(z)}{2}E\left[\frac{D^1}{NTh^4}\right]^{-1} \rightarrow m^2(z)\left(\mu_2 f(z,z)\right)^{-1} + o(1) = R$.

E$\left(\frac{D^2}{NTh^4}\right) = O(h) = o(1)$ and $E\left(\frac{D^3}{NTh^4}\right) = O(h) = o(1)$.

Also, $E(D^4)^2 = \sum\sum E\left(\Delta^2 z_{it}\Delta^2 u_{it}K^2(z_{it})K^2(z_{it-1})\right) = NTh^4 4\sigma^2 f(z,z)(\mu_2 - 1) + O(h^5)$. $E\left(\frac{D^4}{NTh^4}\right)^2 = \frac{4\sigma^2 f(z,z)(\mu_2-1)}{(NTh^4)} + O(h^5)$.

Thus $\sqrt{NTh^4}var\left(\frac{D^4}{NTh4}\right) = \Phi + o(1)$, where $\Phi = 4\sigma^2 f(z,z)(\mu_2 - 1)$.

By Lindberg-Levy Central Limit theorem $\sqrt{NTh^4}\left(\frac{D^4}{NTh^4}\right) \xrightarrow{d} N(0,\Phi)$

Thus it is proved that $\sqrt{NTh^4}\left(\hat{\beta}(z) - \beta(z)\right) \tilde{\ } N(0,\Sigma)$, where $\Sigma = R^{-1}\Phi R^{-1}$

7 Appendix B

For asymptotic normality of $\tilde{\beta}(z)$

$$\tilde{\beta}(z) - \beta(z) = \frac{m^2(z)}{2} \frac{\sum_i \sum_t K_{it}(z_{it} - \bar{z}_{i.})\left[(z_{it} - z)^2 - (\bar{z}_{i.} - z)^2 + (u_{it} - \overline{u}_{i.}\right]}{\sum_i \sum_t (z_{it} - \bar{z}_{i.})^2 K_{it}}$$

First I state some lemmas:

$$E(z_{it}^2 K_{it} \approx hz^2 f(z) + O(h^2) \tag{B.1}$$

$$E(\bar{z}_{i.}^2 K_{it}) \approx hT^2 z f(z) + O(h^2) \tag{B.2}$$

$$E(z_{it} \bar{z}_{i.} K_{it}) \approx hz^2 f(z) + O(h^2) \tag{B.3}$$

$$E(z_{it}^3 K_{it}) \approx hz^3 f(z) + O(h^2) \tag{B.4}$$

$$E(z_{it} \bar{z}_{i.}^2 K_{it}) \approx hz^3 T^2 f(z) + O(h^2) \tag{B.5}$$

$$E(z_{it}^2 \bar{z}_{i.} K_{it}) \approx hz^3 f(z) + O(h^2) \tag{B.6}$$

So using (B1-B3) the expectation of the denominator:

$$\frac{1}{NTh} E(\sum_i \sum_t (z_{it} - \bar{z}_{i.})^2 K_{it}) \approx T^2 z^2 f(z) + o(1) = R_1$$

Similarly using (B1-B6)

$$\frac{1}{NT} \sum_i \sum_t E(K_{it}(z_{it} - \bar{z}_{i.})\left[(z_{it} - z)^2 - (\bar{z}_{i.} - z)^2\right]) \approx o(1)$$

The second moment of $\sum \sum K_{it}(z_{it} - \bar{z}_{i.}) U_{it}$

$$\sqrt{NTh^3} E\left(\frac{\sum \sum K_{it}(z_{it} - \bar{z}_{i.}) U_{it}}{NTh^3}\right)^2 \approx \phi_1 + o(1) \tag{B.7}$$

$$\text{where } \phi_1 = \frac{(T-1)}{T} \sigma_u^2 [\int K^2(\psi_{1t}) \psi_{1t}^2 d\psi_{1t}$$

By Lindberg-Levy Central Limit theorem

$$\sqrt{nTh^3}\left(\sum\sum K_{it}\left(z_{it}-\bar{z}_{i.}\right)U_{it}\right)\xrightarrow{d}N\left(0,\phi_1\right)$$

Thus,

$$\sqrt{NTh}\left(\tilde{\beta}(z)-\beta(z)\right)\tilde{\ }N\left(0,\Sigma_1\right)$$

$$\text{where } \Sigma_1 = R_1^{-1}\Phi_1 R_1^{-1}$$

8 References

Angrist, J. D. and A. B. Krueger (1991), Does Compulsory School Attendance Affect Schooling and Earnings? *Quarterly Journal of Economics*, 106, 979-1014.

Baltagi, B. H. (2000), *Econometric Analysis of Panel Data*, John Wiley, New York.

-(1996), Panel Data Methods, forthcoming in, Handbook of Applied Economic Statistics (A. Ullah and D. E. A. Giles, ed.) Marcel Dekker, New York.

Baltagi, B. H., Y. J. Chang, and Q. Li (1992), Monte Carlo Results on Several New and Existing Tests for the Error Component Model, *Journal of Econometrics*, 54, 95-120.

Berg, M. D., Q. Li and A. Ullah (1999), Instrumnetal Variable Estimation of Semiparametric Dynamic Panle Data Models: Monte Carlo Results on Several New and Existing Estimators, in *Nonstationary Panel, Panel Cointegration and Dynamic Panels* (B.H. Baltagi ed.) JAI Press.

Cai, Z., J. Fan, and Q. Yao (2000), Functional-Coefficient Regression Models for Nonlinear Time Series, *Journal of the American Statistical Association*, 88, 298-308.

Chamberlain, G. (1984), *Panel Data: Handbook of Econometrics* (Z. Griliches and M. Intriligator ed.) North-Holland, Amsterdam, 1247-1318.

Hsiao, C. (2003), *Analysis of Panel Data*, Econometric Society Monographs, Cambridge University Press.

Kniesner T. and Q. Li, (2002), Nonlinearity in Dynamic Adjustment: Semiparametric Estimation of Panel Labor Supply, *Empirical Economics*, 27, 131-48.

Lee, Tae-Hwy and A. Ullah (2003), Nonparametric Bootstrap Specification Testing in Econometric Models, *Computer-Aided Econometrics*, Chapter 15, edited by David Giles, Marcel Dekker, New York, pp. 451-477.

Li Q. and A. Ullah (1998), Estimating Partially Linear panel Data Models with one-way Error Components *Econometric Reviews* 17(2):145-166.

Li, Q. and T. Stengos (1996), Semiparametric Estimation of Partially Linear Regression Model, *Journal of Econometrics* 71:389-397.

Matyas, L. and P. Sevestre (1996), ed., *The Econometrics of Panel Data: A Handbook of the Theory with Applications*, Kluwer Academic Publishers, Dordrecht.

Mundlak, Y. (1978), On the Pooling of Time Series and Cross-Section Data, *Econometrica* 46, 69-85.

Nerlove, M. (1971), A Note on the Error Components Models, *Econometrica*, 39, 383-96.

Pagan, A. and A. Ullah (1999), *Non-Parametric Econometrics*, Cambridge University Press.

Porter, J. R. (1996), Essays in Econometrics, Ph. D Dissertation, MIT.

Rivera-Batiz, F. L. (1999), Undocumented workers in the labor market: An analysis of the earnings of legal and illegal Mexican immigrants in the United States, 12: 91-116.

Robinson, P. M.(1986), Nonparametric Methods in Specification, *The Economic Journal Supplement*, 96, 134-141.

(1988a), Root-N Consistent Semiparametric Regression, *Econometrica* 56:931-954.

(1988b), Semiparametric econometrics: A Survey, *Journal of Applied Econometrics*, 3(1), 35-51.

Sandler, W. (1992), The Effect of Women's Schooling on Fertility, *Economic Letters*, 40, 229-233.

Ullah, A. and N. Roy (1998), Parametric and Nonparametric Panel Data Models, in *Handbook of Applied Economics and Statistics*, edited by A. Ullah and David E. A. Giles, Marcel Dekker.

Vella, F. and M. Verbeek (1998), Whose Wages Do Unions Rise? A Dynamic Model of Unionism and Wage Rate Determination for Young Men, *Journal of Applied Econometrics* 13, 163-183.

Table 1: Root Mean Square Error Difference between the two Parametric Slope Estimators

Panel A: The value in the cell is $Rmse(\widehat{\beta}_{dev}) - Rmse(\widehat{\beta}_{diff})$ for Nerlove data generation

(N,T)	3	6	10	50	100	500
10	1.076	3.866	6.442	10.276	10.507	10.588
50	1.026	3.868	6.439	10.283	10.512	10.590
100	1.036	3.883	6.435	10.285	10.513	10.590
500	1.040	3.876	6.438	10.286	10.512	10.590
1000	1.043	3.879	6.436	10.287	10.513	10.592

Panel B: The value in the cell is $Rmse(\widehat{\beta}_{dev}) - Rmse(\widehat{\beta}_{diff})$ for Berg, Li and Ullah data generation.

(N,T)	3	6	10	50	100	500
10	0.821	1.182	1.682	3.863	5.525	12.773
50	0.350	0.533	0.750	1.742	2.467	5.589
100	0.246	0.370	0.520	1.237	1.765	3.865
500	0.112	0.174	0.225	0.557	0.790	1.763
1000	0.075	0.121	0.166	0.401	0.549	1.816

Table 2: Nonparametric and the Parametric Slope Estimators: First-Differencing
($b2 = 2,\ \rho = 0.8,\ c1 = 0,\ c2 = 2$,α_i not correlated with $\overline{z}_{i.}$)

	$N = 10$								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widehat{\beta}(z)$	-0.420	0.351	0.547	-0.426	0.233	0.485	-0.419	0.164	0.450
$\widehat{\beta}_{diff}$	-0.0517	2.140	2.139	-0.055	1.406	1.407	-0.011	0.986	0.985
$Rmse\widehat{\beta}_{diff} - Rmse\widehat{\beta}(z)$	1.592			0.922			0.535		
	$N = 50$								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widehat{\beta}(z)$	-0.418	0.141	0.441	-0.417	0.103	0.430	-0.411	0.085	0.420
$\widehat{\beta}_{diff}$	-0.009	0.847	0.847	-0.002	0.619	0.618	0.033	0.508	0.509
$Rmse\widehat{\beta}_{diff} - Rmse\widehat{\beta}(z)$	0.406			0.189			0.089		
	$N = 100$								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widehat{\beta}(z)$	-0.422	0.098	0.434	-0.416	0.075	0.423	-0.149	0.057	0.423
$\widehat{\beta}_{diff}$	-0.032	0.591	0.592	0.003	0.449	0.449	-0.014	0.342	0.343
$Rmse\widehat{\beta}_{diff} - Rmse\widehat{\beta}(z)$	0.158			0.026			-0.080		

Table 3: Nonparametric and the Parametric Slope Estimators: Mean Deviation

($b2 = 2,\ \rho = 0.8,\ c1 = 0,\ c2 = 2$,α_i not correlated with $\bar{z}_{i.}$)

	$N = 10$								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widetilde{\beta}(z)$	0.067	1.842	1.842	0.006	1.267	1.267	0.009	0.948	0.947
$\widehat{\beta}_{dev}$	0.081	2.417	2.417	-0.001	1.635	1.634	0.013	1.189	1.189
$Rmse\widehat{\beta}_{dev} - Rmse\widetilde{\beta}(z)$	0.575			0.367			0.241		
	$N = 50$								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widehat{\beta}(z)$	0.026	0.796	0.796	-0.01	0.590	0.590	0.001	0.406	0.405
$\widehat{\beta}_{diff}$	0.030	1.027	1.026	-0.009	0.745	0.744	0.002	0.517	0.517
$Rmse\widehat{\beta}_{dev} - Rmse\widetilde{\beta}(z)$	0.230			0.154			0.111		
	$N = 100$								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widehat{\beta}(z)$	0.036	0.566	0.567	-0.008	0.389	0.389	0.015	0.294	0.294
$\widehat{\beta}_{diff}$	0.037	0.707	0.708	-0.009	0.488	0.487	0.016	0.373	0.373
$Rmse\widehat{\beta}_{dev} - Rmse\widetilde{\beta}(z)$	0.141			0.099			0.079		

Table 4: Nonparametric and the Parametric Slope Estimators: First-Differencing
($b1 = 0.5, b2 = 2,\ \rho = 0.8,\ c2 = 2,\ c1 = 2$ (α_i correlated with $\overline{z}_{i.}$)

	N = 10								
	T = 3			T = 6			T = 10		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widehat{\beta}(z)$	-0.421	0.330	0.535	-0.409	0.225	0.467	-0.423	0.170	0.456
$\widehat{\beta}_{diff}$	-0.022	1.990	1.990	0.048	1.373	1.373	-0.036	1.025	1.025
$Rmse\widehat{\beta}_{diff} - Rmse\widehat{\beta}(z)$	1.454			0.906			0.569		
	N = 50								
	T = 3			T = 6			T = 10		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widehat{\beta}(z)$	-0.425	0.140	0.447	-0.420	0.100	0.432	-0.416	0.083	0.424
$\widehat{\beta}_{diff}$	-0.047	0.839	0.840	-0.020	0.601	0.601	0.007	0.500	0.500
$Rmse\widehat{\beta}_{diff} - Rmse\widehat{\beta}(z)$	0.392			0.169			0.076		
	N = 100								
	T = 3			T = 6			T = 10		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widehat{\beta}(z)$									
$\widehat{\beta}_{diff}$									
$Rmse\widehat{\beta}_{diff} - Rmse\widehat{\beta}(z)$									

Table 5: Nonparametric and the Parametric Slope Estimators: Mean-Deviation

$(b2 = 2,\ \rho = 0.8,\ c2 = 2,\ c1 = 2$ (α_i correlated with $\overline{z}_{i.}$)

	$N = 10$								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widetilde{\beta}(z)$	0.067	1.842	1.842	0.006	1.267	1.267	0.009	0.948	0.947
$\widehat{\beta}_{dev}$	0.081	2.417	2.417	-0.001	1.635	1.634	0.013	1.189	1.189
$Rmse\widehat{\beta}_{dev} - Rmse\widetilde{\beta}(z)$	0.575			0.367			0.242		
	$N = 50$								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widetilde{\beta}(z)$	0.049	0.844	0.845	-0.012	0.544	0.544	-0.029	0.426	0.427
$\widehat{\beta}_{dev}$	0.049	1.068	1.069	-0.013	0.695	0.694	-0.030	0.538	0.539
$Rmse\widehat{\beta}_{dev} - Rmse\widetilde{\beta}(z)$	0.224			0.15			0.112		
	$N = 100$								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widetilde{\beta}(z)$	-0.001	0.568	0.568	-0.008	0.389	0.389	0.000	0.299	0.299
$\widehat{\beta}_{diff}$	-0.001	0.720	0.720	-0.009	0.488	0.487	0.000	0.388	0.387
$Rmse\widehat{\beta}_{dev} - Rmse\widetilde{\beta}(z)$	0.152			0.098			0.088		

Table 6: Nonparametric and the Parametric Slope Estimators with Increased Nonlinearity: First Differencing

($b2 = 4,\ \rho = 0.8,\ c2 = 2,\ c1 = 0$ (α_i not correlated with z_{it})

	$N = 10$								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widehat{\beta}(z)$	-0.438	0.335	0.551	-0.404	0.236	0.468	-0.421	0.174	0.456
$\widehat{\beta}_{diff}$	-0.118	2.033	2.036	0.084	1.431	1.433	-0.026	1.046	1.046
$Rmse\widehat{\beta}_{diff} - Rmse\widehat{\beta}(z)$	1.485			0.965			0.590		
	$N = 50$								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widehat{\beta}(z)$	-0.414	0.149	0.440	-0.413	0.108	0.427	-0.414	0.082	0.422
$\widehat{\beta}_{diff}$	0.015	0.897	0.897	0.023	0.650	0.650	0.016	0.491	0.491
$Rmse\widehat{\beta}_{diff} - Rmse\widehat{\beta}(z)$	0.456			0.223			0.069		
	$N = 100$								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widehat{\beta}(z)$	-0.417	0.099	0.429	-0.414	0.074	0.420	-0.420	0.060	0.424
$\widehat{\beta}_{diff}$	-0.002	0.594	0.593	0.018	0.445	0.445	-0.020	0.358	0.359
$Rmse\widehat{\beta}_{diff} - Rmse\widehat{\beta}(z)$	0.165			0.025			-0.065		

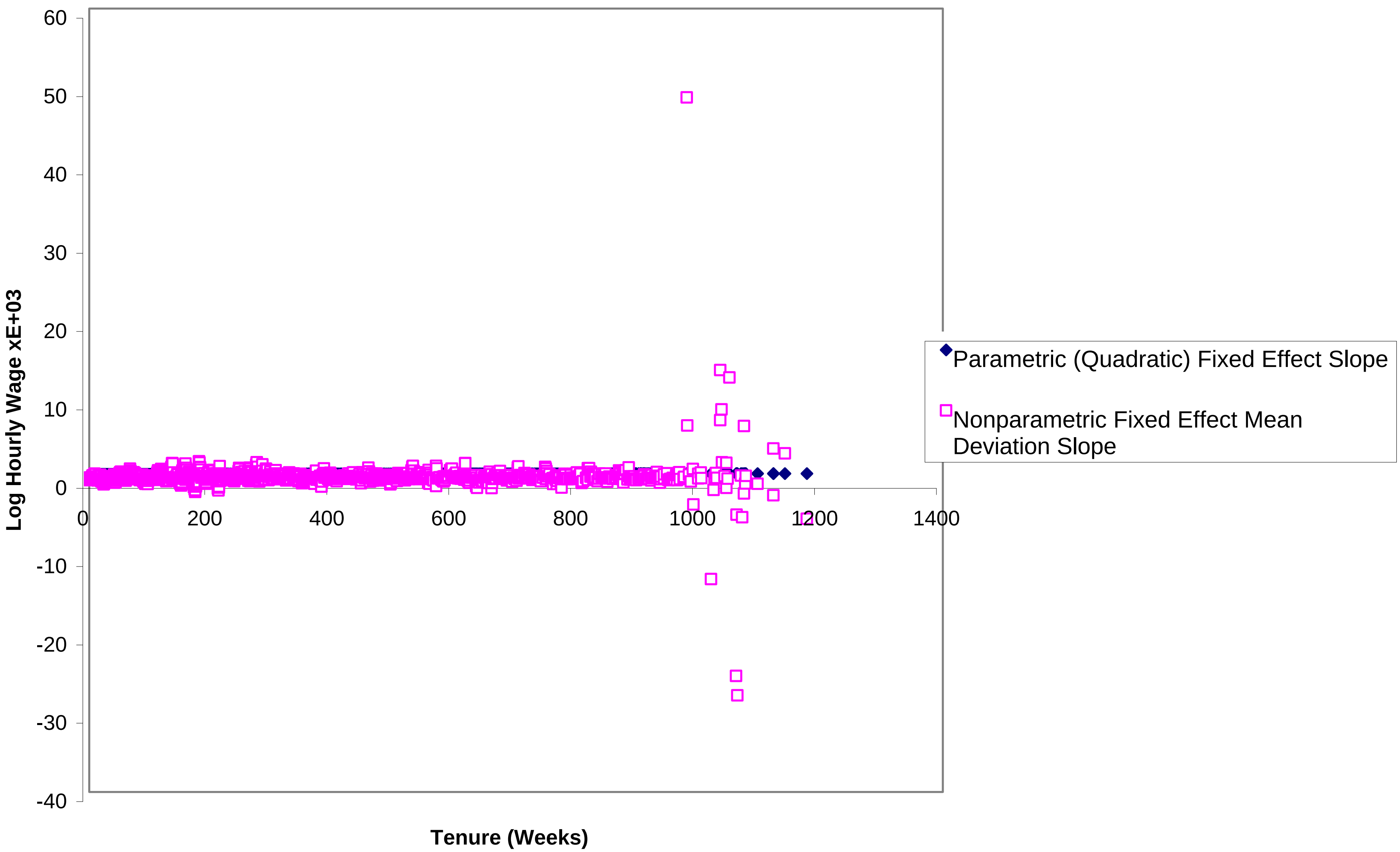
Table 7: Nonparametric and the Parametric Slope Estimators with Increased Nonlinearity: Mean-Deviation

($b2 = 4,\ \rho = 0.8,\ c2 = 2,\ c1 = 0$ (α_i not correlated with $\overline{z}_{i.}$)

	$N = 10$								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widetilde{\beta}(z)$	-0.042	1.985	1.985	0.007	1.283	1.282	0.008	0.958	0.957
$\widehat{\beta}_{dev}$	-0.035	2.578	2.577	0.001	1.654	1.653	0.011	1.202	1.201
$Rmse\widehat{\beta}_{dev} - Rmse\widetilde{\beta}(z)$	0.593			0.371			0.244		
	$N = 50$								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widetilde{\beta}(z)$	0.028	0.804	0.804	-0.011	0.597	0.596	-0.029	0.405	0.406
$\widehat{\beta}_{diff}$	0.032	1.037	1.037	-0.010	0.754	0.753	-0.030	0.511	0.512
$Rmse\widehat{\beta}_{dev} - Rmse\widetilde{\beta}(z)$	0.233			0.157			0.106		
	$N = 100$								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widetilde{\beta}(z)$	-0.030	0.537	0.537	-0.007	0.394	0.394	0.0001	0.301	0.301
$\widehat{\beta}_{diff}$	-0.031	0.672	0.673	-0.008	0.494	0.494	0.0002	0.391	0.390
$Rmse\widehat{\beta}_{dev} - Rmse\widetilde{\beta}(z)$	0.136			0.1			0.09		


Figure 1: Hourly Wage Slope Estimation with respect to Age
Log Hourly wage
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
-0.02
-0.04
-0.06
0
5
10
15
20
25
30
35
40
45
Age (years)
Parametric (Quadratic) Fixed Effect Slope
Nonparametric Fixed Effect Mean Deviation Slope

Figure 2: Hourly Wage Slope Estimation with respect to Tenure



Figre3: Hourly Wage Slope Estimation with respect to Age: First-Differencing

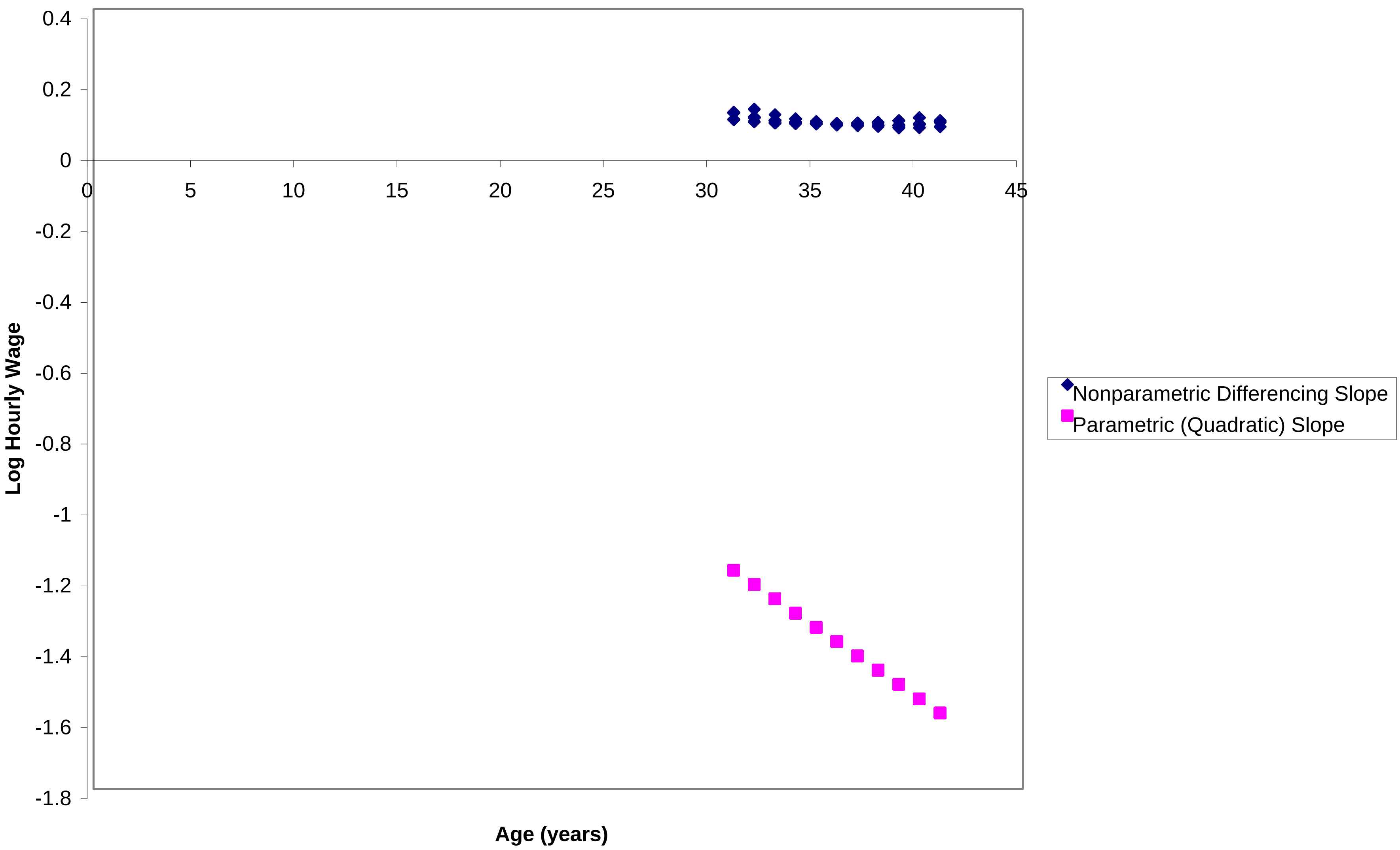


Figure 4: Hourly Wage Slope Estimation with respect to Tenure: First-Differencing