Nonparametric Slope Estimators for Fixed-Effect Panel Data

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

January 2005 (Working Paper)

[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 and at the MEG, Northwestern Oct 2004.

Abstract

In panel data the interest is often in slope estimation while taking account of the unobserved cross sectional heterogeneity. This paper proposes two nonparametric slope estimation where the unobserved effect is treated as fixed across cross section. The first estimator uses first-differencing transformation and the second estimator uses the mean deviation transformation. The asymptotic properties of the two estimators are established and the finite sample Monte Carlo properties of the two estimators are investigated allowing for systematic dependence between the cross-sectional effect and the independent variable. Simulation results suggest that the new nonparametric estimators perform better than the parametric counterparts. We also investigate the finite sample properties of the parametric within and first differencing estimators. A very common practice in estimating earning function is to assume earnings to be quadratic in age and tenure, but that might be misspecified. In this paper we estimate nonparametric slope of age and tenure on earnings using NLSY data and compare it to the parametric (quadratic) effect.

Keywords: Nonparametric, Fixed-effect, Kernel, Monte carlo
JEL Classification: *C1, C14, C23, C15*

1 Introduction

In panel data, we are able to get observations over cross-sectional units over time and can capture the cross-ssectional heterogeneity by including an individual or cross-sectional effect in our model. In traditional panel models whether the cross sectional effect is treated as fixed or random is dependent on how the sample is drawn from the population. 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 is a parameter varying across the cross-section. The cross sectional effect when treated as fixed and non-random, allows it to be correlated with other exogenous regressors in the model.[1] This property of the "fixed" effect is very important in applied econometrics where the cross sectional effect captures omitted variables, allowing them to be correlated with the explanatory variables included in the model. In the contemporary terminology the correlation of the unobserved effect with the independent variable determines whether the cross-sectional time invariant effect is fixed or random, irrespective of whether it is randomly drawn or not.[2] In this paper, we present two nonparametric slope estimation for fixed effect panel models and do monte carlo simulations to investigate the small sample properties of the estimators. Moreover, in the monte carlo simulations we investigate the properties of different estimators allowing for a systematic dependence of the randomly drawn cross sectional effect with the independent variables.

It is well known in theoretical econometrics that misspecification of the functional form leads to biased estimates of the parameters. Often policies are based on these biased estimates, making misspecification of the functional form an important issue in applied econometrics. The significance of the functional form in the econometric modelling makes nonparametric analysis very important. In nonparametric models, no specific functional form is imposed on how the independent variable affects the dependent variable, see Ullah and Pagan (1999). An important extension of non-

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

[2]Accoridng to Wooldridge (2002) the cross-sectional effect at least in microeconometrics is randomly drawn no matter whether it is fixed or random.

parametric kernel techniques have been to panel data models Ullah and Roy (1998), Porter (1996). In particular, there have been extensive work on semiparametric or 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), Berg, Li and Ullah (2000), Ullah and Mundra (2001). However, there has not been enough work on pure nonparametric panel models; recently, Racine and Li (2004) propose a nonparametric estimator with both continuous and count data and Henderson and Ullah (2004) investigate the nonparametric estimation of random effect models. In this paper we propose two nonparametric fixed-effect slope estimation. This is important in applied work where we are often interested in the slope estimate while controlling for the unobserved effect but the estimate of the cross sectional effect is not of interest. For instance, the researcher is interested in the effect of tenure on wages while controlling for the croo-sectional effect, which might represent the ability of the individual. A second examplecan be that a researcher maybe interested in the effect of countrys' income on its international trade while accounting for the country specific effect (capturing infrastructure, institution etc.), one can think of many more examples.

Firstly, in this paper we establish two nonparametric estimators, one using the first-differencing and the second estimator uses within or mean-deviation proposed in Ullah and Roy (1998). We establish the asymptotics of the two estimator and present finite sample monte carlo properties allowing for systematic dependence among the unobserved cross sectional effects, as well as the unobserved effect being correlated with the explanatory variable. We compare the nonparametric estimators with their parametric counterparts. Secondly, this paper uses monte carlo simulations to explore the finite sample properties of the parametric fixed effect estimation: within or mean-deviation and first-differencing. It is well known that asymptotically both the transformations mean-deviation and differencing gives consistent estimates, but little is known how these estimators compare in a finite sample. This paper presents the finite sample properties of the two parametric estimators allowing for the unobserved fixed effect to be random and allowing for the correlation between the cross-sectional effect and the explanatory variable.

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 finite sample results under various forms of dependence between the cross sectional effects and the explanatory variables are discussed in section 4. Section 5 presents the application of the two nonparametric estimates to investigate the effect of age and tenure on earnings using NLSY data. Finally, section 6 concludes.

2 Nonparametric Slope Estimation

The parametric (linear) fixed-effect 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 the exogenous variable and β is the parameter of interest and α_i is the unobserved 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 follow the usual *iid* error structure with mean zero and constant variance.[3] 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 in (2.3) is to estimate the slope $m'(z_{it})$ in (2.2) locally in the interval h around z by linear approximation $(z_{it} - z)\beta(z)$.[4]

[3] In this section we do not impose any well defined cross sectional dependence between α_i and z_{it}, which we do in the monte carlo simulations.

[4] This is similar to the nonparametric kernel regression functional models and 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), and Matyas and Sevestre (1996).[5] In this paper, we use the two transformations within and first-differencing to the fixed effect nonparametric model and estimate the slope coefficients with local linear kernel weighted techniques. In addition, we compare the two methods in the linear parametric model for finite sample properties. For linear models, it is well known that the two methods give consistent estimates and Verbeek (1995) shows that the two transformations gives same estimates for $T = 2$, but finite sample properties of the two slope estimates are unknown. In section 4.1, we explore the monte carlo properties of the two linear parametric slope estimates allowing for the cross-sectional effect to be correlated with the independent variables.

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 local linear estimator of $\beta(z)$ is given by,

$$\hat{\beta}(z) = \sum_{i=1}^{n}\sum_{t=2}^{T} w_{it}\Delta y_{it} \tag{2.5}$$

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.[6]

.

[5]Baltagi and Li (2002) used first-differencing for series estimation of semiparametric panel model.

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

2.2 Deviation from Mean Estimator

Deviation from mean transformation for the panel data model is proposed as follows[7]:

$$\bar{y}_{i.} = \alpha_i + m(z) + (\bar{z}_{i.} - z)\beta(z) + \overline{u}_{i.} \tag{2.6}$$

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 (2.6) from (2.3) gives

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

The local FE estimator of the slope $\beta(z)$ can then be obtained by minimizing $\sum\limits_i\sum\limits_t(y_{it} - \bar{y}_{i.} - (z_{it} - \bar{z}_{i.})\beta(z))^2 k(\frac{z_{it}-z}{h})$.This gives the slope estimator as follows:

$$\tilde{\beta}(z) = \sum_i\sum_t \frac{k_{it}(y_{it} - \bar{y}_{i.})(z_{it} - \bar{z}_{i.})}{\sum\limits_i\sum\limits_t k_{it}(z_{it} - \bar{z}_{i.})^2}, \tag{2.7}$$

2.3 Estimation of the Unobserved Effect

In this paper the parameter of interest is the nonparametric slope $\beta(z)$, the computation procedure of which, similar to the linear parametric panel model does not require the fixed effect to be estimated, Hsiao (2003). If there was an interest in estimating α_i, one can substitute the estimate of $\beta(z)$ from both the deviation and first differencing estimator in $y_{it} - (z_{it} - z)\beta(z) = \alpha_i + m(z) = \delta_i(z)$ and obtain the estimate, $\widehat{\delta}_i(z)$. In order to identify the unobserved cross sectional effect α_i, we will need an additional restriction, lets say $\sum \alpha_i = 0$. This will give the estimate of the nonparametric cross sectional effect $\widehat{\alpha}_i = \widehat{\delta}_i(z) - \frac{\sum \widehat{\delta}_i(z)}{n}$, as the deviation of the nonparametric fixed effect estimator at a point z from the unit (cross-sectional) mean. Note here that $\widehat{m}(z) = \frac{\widehat{\delta}_i(z)}{n}$.[8]

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

[8] This is simlar to the parametric case if we have a fixed effect α_i and a global intercept μ in the model, and we get the estimate of $\alpha_i + \mu = \delta_i$. Together with a simple restriction $\sum \alpha_i = 0$, we identify the cross-sectional effect in the linear panel model as $\widehat{\alpha}_i = \delta_i - \frac{\sum \delta_i}{n}$. The asymptotic properties of the nonparametric cross sectional effect is future research.

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)

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 all Borel measurable, bounded, real valued functions $K(\psi)$ such that (i) $\int K(\psi)d\psi = 1$, (ii) $\int |K(\psi)|\psi d\psi = 0$ (iii) $|\psi||K(\psi| \to 0$ as $|\psi| \to \infty$, (iv) sup $|K(\psi)| < \infty$, (v) $\int K^2(\psi)d\psi < \infty$ (vi) $\int \psi^2 K(\psi)d\psi = \mu_2 < \infty$ (vii) $\int \psi K^2(\psi)d\psi < \infty$ (viii) $\int \psi^2 K^2(\psi)d\psi = \phi_1 < \infty$

Theorem 1: Under the following assumptions

(1) For all t, $(y_{it}\,,\, z_{it})$ are iid. across i and z_{it} is a second order stationary real valued stochastic process $\forall\, i$ 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 $\epsilon\, G_{\mu-1}^2$ for some positive integer $\mu > 2$.

(2) $E(u_{it} \mid z_{it},\, z_{it-1}) = 0\,,\, E(u_{it}^2 \mid z_{it}, z_{it-1}) = \sigma^2 < \infty$ is continuous in z_{it} and z_{it-1}, and $u_{it}\ \forall\, i$ and t.

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

$$\sqrt{NTh^4}\left(\hat{\beta}(z) - \beta(z)\right) \tilde{}\, N\left(0, \Sigma\right)$$

for large N and fixed T, where $R \simeq m^2(z)\left(\mu_2 f(z,z)\right)^{-1}$, $\Phi = 4\sigma^2\mu_2 f\left(z,z\right)\phi_1$, $\Sigma = R^{-1}\Phi R^{-1}$

For the proof of Theorem 1 see Appendix A. The results can be generalized in multivariate context with q elements in z_{it}, replace NTh^4 by NTh^{2q+2}.

Theorem 2: Under the following assumptions

(1) For all t, $(y_{it}\,,\, z_{it})$ are iid. across i and z_{it} is a second order stationary real valued stochastic process $\forall\, i$ and z_{it} admits a density function $g \epsilon G_{\mu-1}^\infty$, $m(z_{it})\, \epsilon\, G_{\mu-1}^2$ for some positive integer $\mu > 2$.

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

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

$$\sqrt{NTh^3}\left(\tilde{\beta}(z)-\beta(z)\right)\tilde{\ }N\left(0,\Sigma_1\right)$$

where $\Sigma_1 = R_1^{-1}\Phi_1 R^{-1}$, where $R_1 = T^2 z^2 g(z)$ and $\Phi_1 = \sigma_u^2 g(z).K_1$, where K_1 is some function of ϕ_1 and $\int \psi K^2(\psi)d\psi$.For the proof of the Theorem 2 see Appendix B. The results can be generalized in multivariate context with q elements in z_{it}, replace NTh^3 by NTh^{q+2}..

4 Monte Carlo Results

In this section we discuss the Monte Carlo properties of the within and first differencing estimator, when the unobserved effect is randomly drawn both for the parametric and the nonparametric models. It is well known that for large N and fixed T , both the deviation and first differencing estimator gives consistent estimate of the slope in panel models. Though in applied work we are often far from large N, and it becomes important to investigate how the estimated slope compares under the two transformations in finite sample. The monte carlo properties of the estimated slope are investigated both when α_i is correlated with z_{it} and when it is not.

4.1 Parametric Models

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. In these experiments z_{it} is generated by the following data generating methods

(i) DGP1: $z_{it} = o.1t + 0.5z_{it-1} + w_{it}$, where $z_{io} = 10 + 5w_{io}$ and $w_{it} \sim U[-0.5, 0.5]$, u_{it} is drawn from standard normal distribution, this mechanism was followed by Baltagi et al.(1992), Li and Ullah (1992) and was first proposed by Nerlove (1971).

(ii) DGP2: $z_{it} \sim U[-\sqrt{3}, \sqrt{3}]$, this DGP was used by Berg et al.(1999).

The model in (4.1) is estimated by both the transformations, deviation from mean and first differencing. The parametric OLS estimator are given as follows:

(1) Parametric first differencing estimator

$$\widehat{\beta}_{diff} = \frac{\sum\sum(z_{it} - z_{it-1})(y_{it} - y_{it-1})}{\sum\sum(z_{it} - z_{it-1})^2} \tag{4.2}$$

(2) Parametric mean deviation estimator

$$\widehat{\beta}_{dev} = \frac{\sum\sum(z_{it} - \overline{z}_{i.})(y_{it} - \overline{y}_{i.})}{\sum\sum(z_{it} - \overline{z}_{i.})^2} \tag{4.3}$$

The results are based on 2000 replications(M) and both N and T are allowed to vary and β is fixed at 8. The number of cross section N takes the values 10, 50, 100, 500, T is varied to be 3, 6, 10, 50,100, and 500. In every experiment we report the Bias, Standard Error, and the Root Mean Square Error for the estimate of the slope β.

$$Bias = M^{-1}\sum_{j=1}^{M}\left(\widehat{\beta}_j - \beta\right),\ RMSE = \left\{M^{-1}\sum_{j=1}^{M}\left(\widehat{\beta}_j - \beta\right)^2\right\}^{-1/2}.$$

The results are given in Table 1 (Panel 1 for DGP1 and Panel 2 for DGP2). From both the Panels 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 mean deviation and the first-differencing estimator is steadily rising as T goes up for fixed N. The bias and the standard error for the first differencing estimator is lower than the mean-deviation estimator as T increases for all N.[9] On the contrary for fixed T, increasing N, in the case of DGP1 there is no significant change in the magnitude of rmse for the two estimates. For DGP2 on the other hand for fixed T and increasing N, deviation is doing better than differencing.

In another experiment for DGP2 we generate α_i as a random variable drawn from v_i, where $v_i \tilde{} N(0, \sigma_v)$, the value of $\sigma_v^2 + \sigma_u^2 = 20$ and $\rho = \sigma_v^2/(\sigma_v^2 + \sigma_u^2)$ takes the value of 0.8. In Table 2 (Panel 1) gives the difference of the root mean square error

[9] According to Verbeek (1995), the two transformations mean-deviation or within and the first differencing gives same results when $T = 2$. For $T >= 2$, if after differencing transformation we keep the time period as T and not $T-1$ for every i, (in other words we keep the redundant variable $z_{i1} - z_{i0}$) then OLS (within) and first-differencing gives the same estimate. In these monte carlo experiments after differencing T is becoming $T - 1$, in this case within OLS is only same as differencing GLS in small samples.

between the within and the first differencing estimator. In Panel 2 in Table 2 we present results from another experiment where α_i is random but also correlated with $\overline{z}_i$. From Table 2 we see that the difference in rmse for the two estimators is similar in both the panels. In both the panels we see that similar to Table 1 the differencing is doing better than deviation as T increases $\forall$ N. Moreover, for fixed T increasing N deviation is doing better than differencing.

4.2 Nonparametric Models

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

$$y_{it} = \alpha_i + z_{it}\beta_1 + z_{it}^2\beta_2 + u_{it} \tag{4.4}$$

where $z_{it} \sim U[-0.5, 0.5]$ by DGP2 and α_i is generated by $\alpha_i = v_i + c_1\alpha_j$, 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. The value of $\sigma_v^2 + \sigma_u^2 = 20$ and $\rho = \sigma_v^2/(\sigma_v^2 + \sigma_u^2)$ 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 and $c_1 = 0$ or $c_1 = 2$. When $c_1 = 0$, we do not allow for any correlation between α_i and α_j, but when $c_1 = 2$, we are allowing for α_i to be correlated with α_j (in some AR fashion). Note that under both situation, the two transformations within and first-differencing will eliminate the unobserved effect and the estimate of the slope will not be effected. 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. Table 3 (Table 5) and Table 4 (Table 6) presents the differencing transformation (mean deviation) for $c_1 = 0$ and $c_1 = 2$ respectively. We see that the nonparametric estimator is consistent and performs better than the parametric estimator for all the cases. For fixed N and increasing T (also for fixed T and increasing N) for both the estimators the difference in the rmse is falling between the parametric and the nonparametric estimators in all the cases. We see that the difference in the rmse for the parametric and the nonparametric estimator falls when α_i is allowed to be correlated with α_j. Compared to first differencing transformation for the mean deviation case the difference between the nonparametric and the parametric rmse is lower.

In another exercise, we allow α_i to be correlated with $\overline{z}_{i.}$ by $\alpha_i = v_i + c_1\alpha_j + c_2\overline{z}_{i.}$, where the value of $c_2 = 0.5$ and c_1 takes the value 0 or 2 (i.e. both when the unobserved effect α_i is not allowed to be correlated with α_j and when it is). The results from the simulation are given in Table 7 (Table 9) and Table 8 (Table 10) for first differencing (mean deviation) for $c_1 = 0$ and $c_1 = 2$ respectively. Here again we see that the nonparametric estimator is doing better than the parametric and for both the estimators for fixed N and increasing T (also for fixed T and increasing N) the difference in the rmse is falling between the parametric and the nonparametric estimators. Moreover, the difference in the rmse for the parametric and the nonparametric estimator falls when α_i is allowed to be correlated with α_j. Also, compared to first differencing transformation for the mean deviation case the difference between the nonparametric and the parametric rmse is lower.

In another experiment we increased the degree of correlation between the random cross sectional effect and the independent variable. In Table 11 we present results from experiment where $\alpha_i = v_i + c_1\alpha_j + c_2\overline{z}$, $c_1 = 0$ and $c_2 = 4$, for first differencing and Table 12 shows for mean deviation. Comparing to Table 8 (where $c_1 = 0$ and $c_2 = 2$) in Table 11 we see that for $N = 10$ and any T, the parametric estimator is doing worse, the difference between the nonparametric and parametric estimator increases. Similarly is the case with mean deviation. So we find evidence that when we increase the correlation between the random cross-sectional effect and the independent variable, the misspecified parametric model performs worse than the nonparametric model.

We also increased the degree of nonlinearity in the model given by (4.4), by increasing the value of β_2 from 2 to 4 and $\alpha_i = v_i + c_1\alpha_j + c_2\overline{z}_{i.}$, where c_2 =0.5 and $c_1 = 2$. From Table 13 (compared to Table 7) for first differencing and Table 14 (compared to Table 9) for mean deviation we see that in small samples the nonparametric estimator is doing better than the parametric estimator; as expected.

5 Application

In this part, we apply the nonparametric estimators to investigate the effect of worker's age and tenure on their earnings 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 time points.[10] In estimating earning functions it is a very common practice to assume that workers earnings are quadratic in age and tenure, see Angrist and Krueger (1991), Sander (1992), Vella and Verbeek (1998) and Rivera-Batiz (1999) to name a few. In the nonparametric model no functional form is imposed on the effect of age and tenure on earnings.

Worker earnings are measured in hourly wages, age in years, and tenure in number of weeks. A parametric fixed-effect (quadratic) model is fit to investigate the effect of age and tenure on workers log hourly wages for a sample of 1000 individuals for $t = 3$ (the years are 1994, 1996, and 1998). The parametric model is estimated by both the first-differencing and the mean-deviation methods and the slope estimates are given in 4.2 and 4.3. Similarly the nonparametric first-differencing and mean-deviation slopes are estimated, given in 2.5 and 2.7. The slope estimates are used to calculate earning elasticity with respect to age and tenure.

Figure 1 and Figure 2 gives the wage elasticity with respect to age and tenure respectively, by first-differencing methods both for the parametric and nonparametric models. Figure 3 and Figure 4 give the same for the mean-deviation method. From Figure 1 we see that nonparametric wage elasticity with respect to age lies mostly between 2 and 3, whereas the parametric first-differncing elasticity is between 0 and -1.5. From Figure 1 and Figure 3, we see that the parametric wage elasticity with increasing age is falling both for the differencing and mean-deviation transformation. For nonparametric wage elasticity we find that the range is bigger and the magnitude in mean-deviation is lower than the first-differencing. Figure 2 shows that the wage elasticity is steadily rising with tenure in the first-differencing parametric case, whereas in the nonparametric case we see that the earning elasticity is rising but at

[10] 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.

an increasing rate.[11] From Figure 4 we see that the parametric mean-deviation wage elasticity with respect to tenure is mostly zero but in the nonparametric case we see high wage elasticity for some workers at a higher level of tenure.

6 Conclusion

The two nonparametric slope estimator proposed in this paper for fixed-effect panel model performs better than the parametric counterparts. Moreoevr, the nonparametric estimator performs better than the parametric estimator under various scenarios of systematic dependence among the random cross sectional effects and also when a correlation is introuced between the random cross-sectional effect and the independent variables in the model. We also find that for the linear fixed-effect estimator, the rmse for the first-differencing estimator is lower than the mean-deviation as T is rising. A simple application of the two nonparametric slope estimator to the NLSY sample exploring the earning elasticity with respect to worker age and tenure shows that the nonparametric results are very different from the parametric, both in the magnitude and the change of the slope.

[11] This might be the case because the workers in the sample are relatively young mostly between the ages 31 - 35.

7 Appendix

7.1 Proof of Theorem 1

For $q=1$, $\hat{\beta}(z)=\sum_{i=1}^{n}\sum_{t=2}^{T}w_{it}\Delta y_{it}$.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 to higher q. Write, $E(\hat{\beta}(z)/z_{it},\ z_{it-1}) = E(\sum\sum w_{it}\left(m(z_{it})-m(z_{it-1})\right))$.

. 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.

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

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

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

$$E(D^4)=E[\sum_i\sum_t\Delta z_{it}\Delta u_{it}K_{it}K_{it-1}] == 0 \tag{A.4}$$

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

$f_{01}[x,y]$ 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.

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

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

Also, $E(D^4)^2 = \sum\sum E\left(\Delta^2 z_{it}\Delta^2 u_{it}K^2\left(z_{it}\right)K^2\left(z_{it-1}\right)\right) = NTh^4 4\sigma^2 f(z,z)\phi_1 + O(h^5)$. $E\left(\frac{D^4}{NTh^4}\right)^2 = \frac{4\sigma^2 f(z,z)\phi_1}{(NTh^4)} + O(h^5)$.

Thus $\sqrt{NTh^4}var\left(\frac{D^4}{NTh4}\right) = \Phi + o(1)$, where $\Phi = 4\sigma^2 f\left(z,z\right)\phi_1$.

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

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

8 Appendix B

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

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

First we state some lemmas:

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

$$E(\bar{z}_{i.}^2 k_{it}) \approx hz^2 g(z) + O(h^2) \tag{B.2}$$

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

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

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

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

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

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

Similarly using (B1-B6)

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

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

$$\sqrt{NTh^3} E\left(\frac{\sum\sum k_{it}\left(z_{it} - \bar{z}_{i.}\right) U_{it}}{NTh^3}\right)^2 \approx \Phi_1 + o(1) \tag{B.7}$$

where $\Phi_1 = \sigma_u^2 g(z).K_1$,where $K_1$is some function of $\phi_1$and $\int \psi K^2(\psi) d\psi$

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,

$$\sqrt{NTh^3}\left(\tilde{\beta}(z) - \beta(z)\right) \tilde{} N\left(0, \Sigma_1\right)$$

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

9 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.

Baltagi, B. and S. Li (2002), Series Estimation of Partially Linear panel Data Models with Fixed Effects, in *Annals of Economics and Finance* 3, 103 - 116.

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.

Henderson, D. and A. Ullah (2004), Nonparametric Random Effect Estimator, Working Paper, University of California, Riverside.

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, *Journal of Population Economics*,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.

Racine, J. and Q. Li (2004), Nonparametric Estimation of Regression Functions with Both Categorical and Continuous Data, *Journal of Econometrics* 119, 99 - 130.

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.

Verbeek, Marno (1995), Alternative Transformations to Eliminate Fixed Effects, *Econometric Reviews*, 14(2), 205 - 211.

Wooldridge, J. (2002), *Econometric Analysis of Cross Section and Panel Data*, Cambridge, MA: MIT Press.

Table 1: Root Mean Square Error Difference between the two Parametric Slope Estimators. The Value in the cell is $Rmse(\widehat{\beta}_{dev}) - Rmse(\widehat{\beta}_{diff})$

Panel (1): DGP1

(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 (2): DGP2

(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: Root Mean Square Error Difference between the two Parametric Slope Estimators when α_i is random. The Value in the cell is $Rmse(\widehat{\beta}_{dev}) - Rmse(\widehat{\beta}_{diff})$.

Panel 1: DGP2 when α_i is not correlated with $\overline{z_i}$

(N,T)	3	6	10	50	100	500
10	-0.159	0.009	-0.048	-0.037	-0.024	-0.012
50	0.014	-0.030	-0.026	-0.016	-0.013	-0.005
100	0.031	-0.009	-0.018	-0.011	-0.008	-0.004
500	0.005	-0.006	-0.009	-0.006	-0.004	-0.002
1000	0.006	-0.005	-0.006	-0.004	-0.003	-0.001

Panel 2: DGP2 when α_i is correlated with $\overline{z_i}$

(N,T)	3	6	10	50	100	500
10	-0.159	0.009	-0.048	-0.037	-0.027	-0.015
50	0.014	-0.030	-0.026	-0.020	-0.013	-0.006
100	0.032	-0.017	-0.018	-0.011	-0.008	-0.004
500	0.005	-0.003	-0.008	-0.006	-0.004	-0.002
1000	0.006	-0.005	-0.006	-0.004	-0.003	-0.001

Table 3: Nonparametric and the Parametric Slope Estimators:
First-Differencing $(b1 = 0.5,\ b2 = 2,\ \rho = 0.8,\ c1 = 0)$

	N=10								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widehat{\beta}(z)$	-0.416	0.420	0.591	-0.402	0.274	0.487	-0.414	0.220	0.469
$\widehat{\beta}_{diff}$	0.013	2.526	2.525	0.089	1.671	1.673	0.019	1.328	1.328
$\text{Rmse}\widehat{\beta}_{diff} - \text{Rmse}\widehat{\beta}(z)$	1.933			1.186			0.859		
	$N = 50$								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widehat{\beta}(z)$	-0.426	0.175	0.460	-0.419	0.119	0.435	-0.412	0.096	0.423
$\widehat{\beta}_{diff}$	-0.055	1.050	1.051	-0.012	0.713	0.713	0.027	0.575	0.575
$Rmse\widehat{\beta}_{diff} - Rmse\widehat{\beta}(z)$	0.406			0.277			0.152		
	$N = 100$								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widehat{\beta}(z)$	-0.424	0.120	0.441	-0.415	0.094	0.425	-0.419	0.069	0.425
$\widehat{\beta}_{diff}$	-0.045	0.721	0.722	0.012	0.563	0.563	-0.014	0.415	0.415
$Rmse\widehat{\beta}_{diff} - Rmse\widehat{\beta}(z)$	0.281			0.138			-0.010		

Table 4: Nonparametric and the Parametric Slope Estimators: First-Differencing ($b1 = 0.5,\ b2 = 2,\ \rho = 0.8,\ c1 = 2$)

	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
$\mathrm{Rmse}\widehat{\beta}_{diff}-\mathrm{Rmse}\widehat{\beta}(z)$	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 5: Nonparametric and the Parametric Slope Estimators:
Mean Deviation $(b1 = 0.5,\ b2 = 2,\ \rho = 0.8,\ c1 = 0)$

	N=10								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widehat{\beta}(z)$	-0.096	2.242	2.242	-0.025	1.061	1.061	0.060	0.925	0.927
$\widehat{\beta}_{diff}$	-0.095	2.943	2.943	-0.026	1.370	1.369	0.063	1.186	1.187
$\mathrm{Rmse}\widehat{\beta}_{diff} - \mathrm{Rmse}\widehat{\beta}(z)$	0.701			0.308			0.26		
	$N = 50$								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widehat{\beta}(z)$	0.038	0.926	0.926	0.002	0.578	0.577	-0.003	0.407	0.406
$\widehat{\beta}_{diff}$	0.042	1.200	1.200	0.003	0.725	0.725	-0.002	0.518	0.518
$Rmse\widehat{\beta}_{diff} - Rmse\widehat{\beta}(z)$	0.274			0.148			0.112		
	$N = 100$								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widehat{\beta}(z)$	0.014	0.585	0.585	-0.013	0.401	0.401	0.003	0.283	0.283
$\widehat{\beta}_{diff}$	0.012	0.732	0.732	-0.013	0.506	0.506	0.004	0.357	0.357
$Rmse\widehat{\beta}_{diff} - Rmse\widehat{\beta}(z)$	0.147			0.105			0.074		

Table 6: Nonparametric and the Parametric Slope Estimators:
Mean Deviation $(b1 = 0.5,\ b2 = 2,\ \rho = 0.8,\ c1 = 2)$

	$N = 10$								
	$T = 3$			$T = 6$			$T = 10$		
	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 7: Nonparametric and the Parametric Slope Estimators:
First-Differencing $(b1 = 0.5, b2 = 2, \rho = 0.8, c1 = 0, c2 = 2, \alpha_i$ correlated with $\overline{z}_{i.})$

	$N = 10$								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widehat{\beta}(z)$	-0.418	0.379	0.564	-0.421	0.278	0.504	-0.424	0.192	0.465
$\widehat{\beta}_{diff}$	0.004	2.314	2.313	-0.028	1.683	1.682	-0.042	1.152	1.152
$Rmse\widehat{\beta}_{diff} - Rmse\widehat{\beta}(z)$	1.749			1.178			0.687		
	$N = 50$								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widehat{\beta}(z)$	-0.426	0.175	0.460	-0.419	0.119	0.435	-0.412	0.096	0.423
$\widehat{\beta}_{diff}$	-0.055	1.050	1.051	-0.012	0.601	0.713	0.027	0.575	0.575
$Rmse\widehat{\beta}_{diff} - Rmse\widehat{\beta}(z)$	0.590			0.277			0.152		
	$N = 100$								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widehat{\beta}(z)$	-0.424	0.120	0.441	-0.413	0.084	0.422	-0.419	0.072	0.425
$\widehat{\beta}_{diff}$	-0.045	0.721	0.722	0.019	0.503	0.503	-0.016	0.431	0.431
$Rmse\widehat{\beta}_{diff} - Rmse\widehat{\beta}(z)$	0.281			0.082			0.005		

Table 8: Nonparametric and the Parametric Slope Estimators: First-Differencing

($b1 = 0.5, b2 = 2, \rho = 0.8, c1 = 2, c2 = 2, \alpha_i$ correlated with $\bar{z}_{i.}$)

	$N = 10$								
	$T = 3$			$T = 6$			$T = 10$		
	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
$\widehat{\beta}(z)$	-0.423	0.098	0.434	-0.412	0.072	0.418	-0.418	0.054	0.422
$\widehat{\beta}_{diff}$	-0.036	0.591	0.592	0.028	0.0434	0.435	-0.009	0.324	0.323
$Rmse\widehat{\beta}_{diff} - Rmse\widehat{\beta}(z)$	0.158			0.016			-0.098		

Table 9: Nonparametric and the Parametric Slope Estimators: Mean-Deviation

($b2 = 2,\ \rho = 0.8,\ c1 = 0,\ c2 = 2,\ \alpha_i$ correlated with $\overline{z_{i.}}$

	N = 10								
	T = 3			T = 6			T = 10		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widetilde{\beta}(z)$	-0.096	2.242	2.242	-0.025	1.061	1.061	0.060	0.925	0.927
$\widehat{\beta}_{dev}$	-0.095	2.943	2.943	-0.026	1.370	1.369	0.063	1.186	1.187
$Rmse\widehat{\beta}_{dev} - Rmse\widetilde{\beta}(z)$	0.701			0.305			0.260		
	N = 50								
	T = 3			T = 6			T = 10		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widetilde{\beta}(z)$	0.038	0.926	0.926	0.002	0.578	0.577	-0.003	0.407	0.406
$\widehat{\beta}_{dev}$	0.042	1.200	1.200	0.003	0.725	0.725	-0.002	0.518	0.518
$Rmse\widehat{\beta}_{dev} - Rmse\widetilde{\beta}(z)$	0.274			0.148			0.112		
	N = 100								
	T = 3			T = 6			T = 10		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widetilde{\beta}(z)$	0.014	0.585	0.585	-0.013	0.401	0.401	0.014	0.285	0.285
$\widehat{\beta}_{diff}$	0.012	0.732	0.732	-0.013	0.506	0.506	0.013	0.361	0.361
$Rmse\widehat{\beta}_{dev} - Rmse\widetilde{\beta}(z)$	0.147			0.105			0.076		

Table 10: Nonparametric and the Parametric Slope Estimators:Mean-Deviation

($b2 = 2,\ \rho = 0.8,\ c1 = 2,\ c2 = 2,\ \alpha_i$ 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 11: Nonparametric and the Parametric Slope Estimators with Increased Correlation with $\overline{z_i}$: First Differencing

$(b2 = 2,\ \rho = 0.8,\ c1 = 0,\ c2 = 4)$

	$N = 10$								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widehat{\beta}(z)$	-0.416	0.420	0.591	-0.402	0.274	0.487	-0.414	0.220	0.469
$\widehat{\beta}_{diff}$	0.013	2.526	2.525	0.089	1.671	1.673	0.019	1.328	1.328
$Rmse\widehat{\beta}_{diff} - Rmse\widehat{\beta}(z)$	1.933			1.186			0.859		
	$N = 50$								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widehat{\beta}(z)$	-0.426	0.175	0.460	-0.419	0.119	0.435	-0.412	0.096	0.423
$\widehat{\beta}_{diff}$	-0.055	1.050	1.051	-0.012	0.713	0.713	0.027	0.575	0.575
$Rmse\widehat{\beta}_{diff} - Rmse\widehat{\beta}(z)$	0.590			0.277			0.152		
	$N = 100$								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widehat{\beta}(z)$	-0.424	0.120	0.441	-0.415	0.094	0.425	-0.420	0.071	0.426
$\widehat{\beta}_{diff}$	-0.045	0.721	0.722	0.012	0.563	0.563	-0.018	0.423	0.424
$Rmse\widehat{\beta}_{diff} - Rmse\widehat{\beta}(z)$	0.281			0.138			-0.002		

Table 12: Nonparametric and the Parametric Slope Estimators with Increased Correlation with mean $\overline{z}_i$: Mean-Deviation

$(b2 = 2,\ \rho = 0.8,\ c1 = 0,\ c2 = 4)$

	$N = 10$								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widetilde{\beta}(z)$	0.093	2.131	2.132	-0.069	1.456	1.457	0.023	0.952	0.952
$\widehat{\beta}_{dev}$	0.100	2.780	2.781	-0.072	1.926	1.926	0.023	1.199	1.199
$Rmse\widehat{\beta}_{dev} - Rmse\widetilde{\beta}(z)$	0.649			0.469			0.247		
	$N = 50$								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widetilde{\beta}(z)$	0.038	0.926	0.926	0.002	0.578	0.577	-0.003	0.407	0.406
$\widehat{\beta}_{diff}$	0.042	1.200	1.200	0.003	0.725	0.725	-0.002	0.518	0.518
$Rmse\widehat{\beta}_{dev} - Rmse\widetilde{\beta}(z)$	0.274			0.148			0.112		
	$N = 100$								
	$T = 3$			$T = 6$			$T = 10$		
	Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
$\widetilde{\beta}(z)$	0.014	0.585	0.585	-0.013	0.401	0.401	0.009	0.290	0.290
$\widehat{\beta}_{diff}$	0.012	0.732	0.732	-0.013	0.506	0.506	0.010	0.363	0.363
$Rmse\widehat{\beta}_{dev} - Rmse\widetilde{\beta}(z)$	0.147			0.105			0.073		

Table 13: Nonparametric and the Parametric Slope Estimators with Increased Nonlinearity: First Differencing

$(b2 = 4,\ \rho = 0.8,\ c1 = 2,\ c2 = 0.5)$

	$N = 10$								
	$T = 3$			$T = 6$			$T = 10$		
	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 14: Nonparametric and the Parametric Slope Estimators with Increased Nonlinearity: Mean-Deviation

$(b2 = 4,\ \rho = 0.8,\ c1 = 2,\ c2 = 0.5)$

	$N = 10$								
	$T = 3$			$T = 6$			$T = 10$		
	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: Elasticity of Hourly Wage with respect to Age: First-Differencing

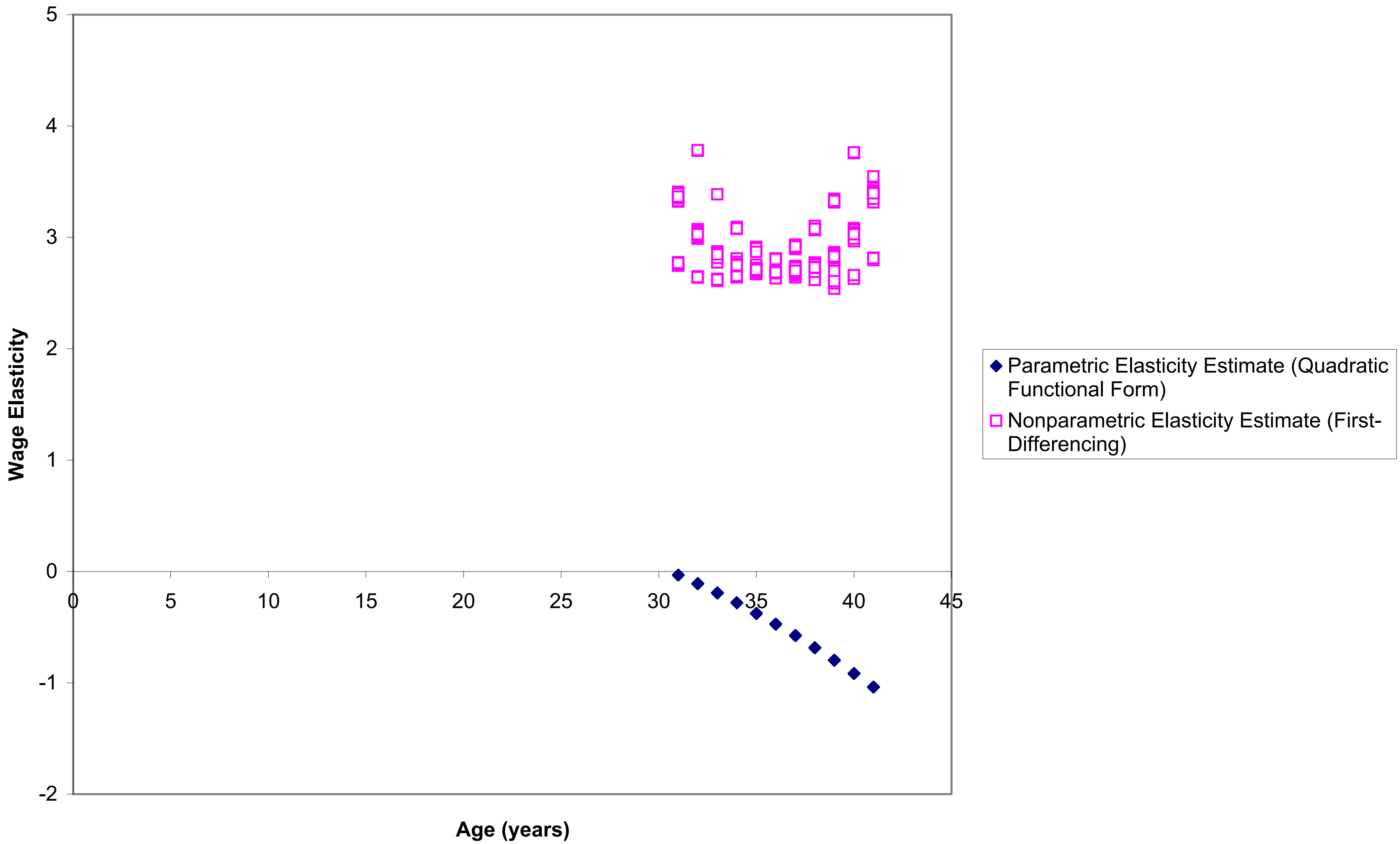

Figure 2: Elasticity of Hourly wage with respect to Tenure: First-Differencing

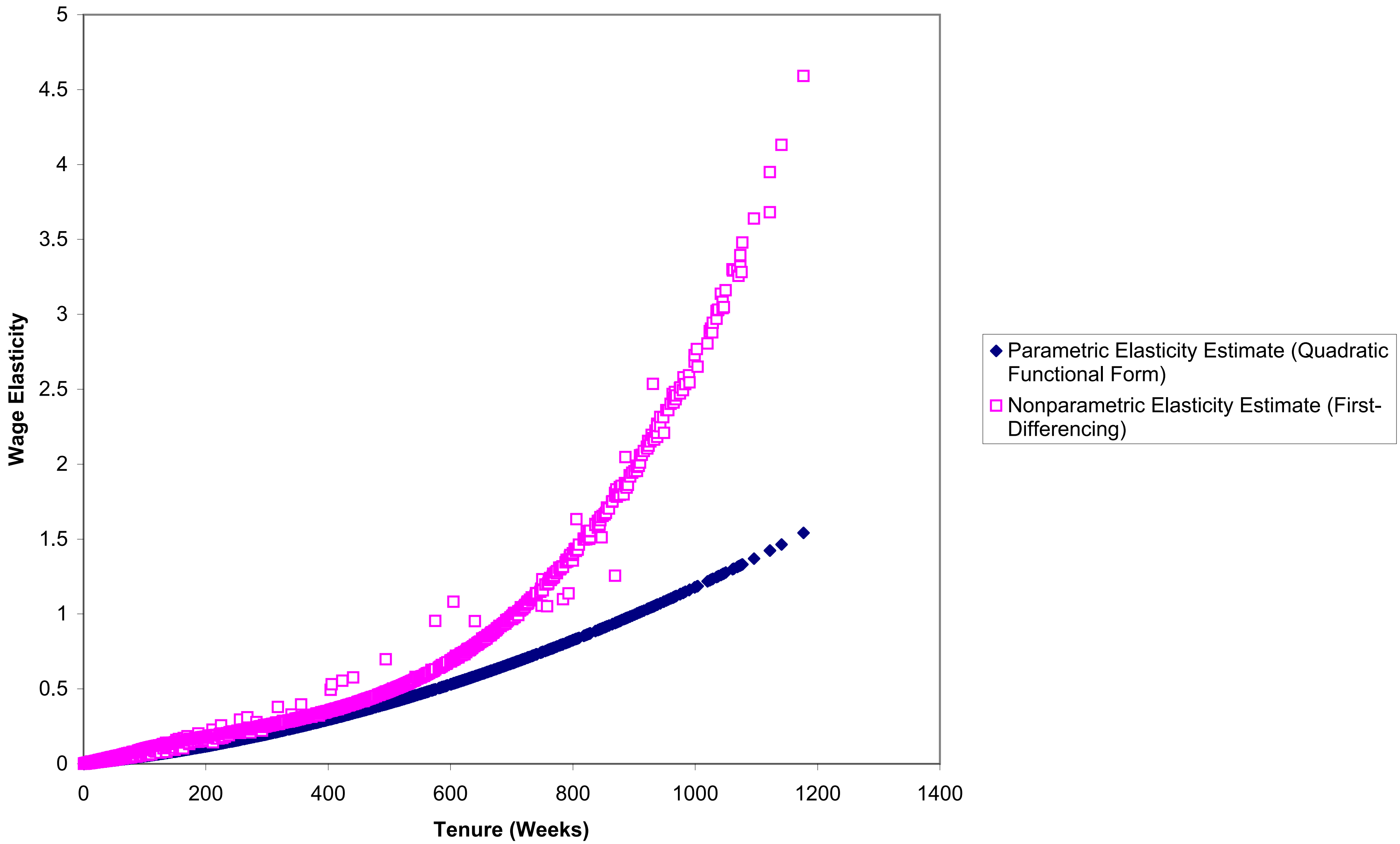

Figure 3 Elasticity of Hourly Wage with respect to Age: Mean-Deviation

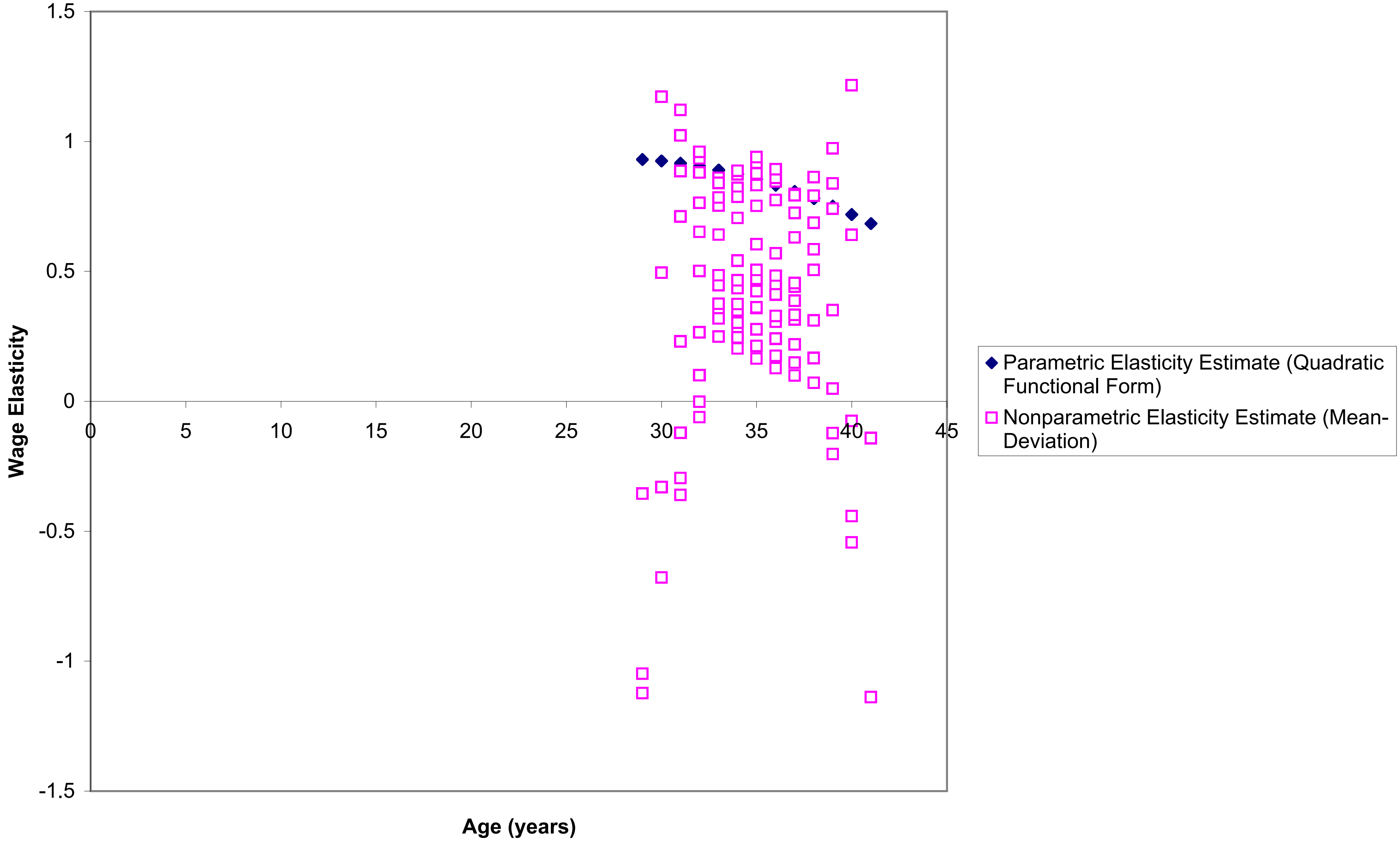

Figure 4: Elasticity of Hourly Wage with respect to Tenure: Mean-Deviation