Analysis of gas phase composition in fuel and exhaust gas in Solid Oxide Fuel Cells

Wei Zhang

Follow this and additional works at: https://researchrepository.wvu.edu/etd

Recommended Citation
https://researchrepository.wvu.edu/etd/7029

This Dissertation is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Dissertation in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/or on the work itself. This Dissertation has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.
Three Essays on Applied Semiparametric Methods

Fan Zhang

Dissertation submitted
to the College of Business and Economics
at West Virginia University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Economics

Feng Yao, Ph.D., Chair
Stratford Douglas, Ph.D.
Eric Olson, Ph.D.
Adam Nowak, Ph.D.
Xiaoli Etienne, Ph.D.

Department of Economics

Morgantown, West Virginia
2017

Keywords: Semiparametric Smooth Coefficient Model; Technical Efficiency;
Economic Freedom; Conditional Factor Model
Copyright 2017, Fan Zhang
ABSTRACT

Three Essays on Applied Semiparametric Methods

Fan Zhang

The first chapter investigates the semiparametric smooth coefficient stochastic production frontier model for panel data in which the distribution of the composite error term is assumed to be of known form but depends on some environmental variables. It proposes multi-step estimators for the smooth coefficient functions as well as the parameters of the distribution of the composite error term and obtain their asymptotic properties. The Monte Carlo study demonstrates that the proposed estimators perform well in finite samples. It also considers an application and perform model specification test, construct confidence intervals, and estimate efficiency scores that depend on some environmental variables. The application uses country level data from the Penn World Table for 134 countries from 1990 to 2011. Results show that two popular parametric models used in the stochastic frontier literature are likely to be misspecified. It also finds that the average efficiency level from the smooth coefficient model are higher than those obtained from the parametric frontier models.

The second Chapter applies a multi-step semiparametric stochastic production frontier estimator proposed in the first chapter to investigate the effects of economic freedom on the production frontier and technical efficiency. It allows output elasticities and technical efficiency to depend on the economic freedom variable, estimate a smooth coefficient stochastic production frontier, and compare with parametric alternatives, the Cobb-Douglas and translog estimates. The results add to the literature on economic freedom and growth in two ways. First, the results highlight the importance of semiparametric approaches as it finds the commonly used parametric approaches restrictive to estimate the marginal productivity of inputs. Second, it finds that the output elasticities of labor, human capital, and physical capital vary with the level of economic freedom.

The third chapter applies a smooth varying coefficient model proposed by Li et al. (2002) to investigate the effects of real interest rate on Fama–French five factor model. This chapter contrasts the conditional factor model with it’s traditional factor model setups. Allowing the abnormal returns and factor loadings to be a smooth function of real interest rate, it observes significant variation on abnormal returns and factor loadings in all the portfolios investigated when real interest rate changes.
## Contents

1 Semiparametric Smooth Coefficient Stochastic Production Frontier ........................................... 1
   1.1 Introduction .................................................................................................................. 1
   1.2 A multi-step semiparametric estimation ................................................................. 3
   1.3 Asymptotic Characterization ...................................................................................... 6
   1.4 Monte Carlo Study ...................................................................................................... 11
      1.4.1 Estimation Procedure ...................................................................................... 11
      1.4.2 Testing Procedure ......................................................................................... 13
   1.5 An Empirical Application ......................................................................................... 15
   1.6 Conclusion ................................................................................................................ 19
   1.7 Appendix 1: Tables and Figures ............................................................................... 21
   1.8 Appendix 2: Proofs .................................................................................................. 32

2 Does Economic Freedom Affect The Production Frontier?
   A Semiparametric Approach With Panel Data ..................................................................... 54
   2.1 Introduction ................................................................................................................ 54
   2.2 Methodology and Model Specification ....................................................................... 56
   2.3 Data Descriptions ...................................................................................................... 59
   2.4 Estimation Results ..................................................................................................... 60
   2.5 Conclusion ................................................................................................................ 63
   2.6 Appendix: Tables and Figures .................................................................................. 64

3 Conditional Factor Model: A Semiparametic Approach ......................................................... 71
   3.1 Introduction ................................................................................................................ 71
   3.2 Methodology and Model Specification ....................................................................... 72
   3.3 Data Description ........................................................................................................ 73
   3.4 Empirical Results ...................................................................................................... 74
   3.5 Conclusion ................................................................................................................ 76
   3.6 Appendix: Tables and Figures .................................................................................. 77
List of Figures

1.1 Plots of smooth coefficient functions against $z_{it}$, where $\alpha_1(z_{it}) = -\frac{1}{1+3z_{it}}$, $\beta_{11}(z_{it}) = (2z_{it})^3$, $\beta_{21}(z_{it}) = ln(5z_{it})$, $\alpha_2(z_{it}) = -5cos(4z_{it})$, $\beta_{12}(z_{it}) = sin(4z_{it})$ and $\beta_{22}(z_{it}) = ln(\frac{z_{it}}{1-z_{it}})$. ................................................................. 29

1.2 Plots of smooth coefficient frontiers and kernel density estimates: left-$\alpha(H)$, middle-$\beta_{L}(H)$, right-$\beta_{K}(H)$. Upper panel: A-smooth coefficient estimates, D and E-95% confidence bound, B-Cobb-Douglas estimate, C-translog estimate. Middle panel: similar to upper, except with 99% confidence bound. Lower panel: kernel density estimates. ................. 30

1.3 Plots of Returns to Scale and kernel density estimates ................................................. 31

1.4 Kernel density estimates of technical efficiency ............................................................. 31

2.1 Plots of Smooth Coefficient Frontiers ............................................................................. 67

2.2 Kernel Density of Composite Error and Technical Efficiency ........................................... 67

3.1 Conditional Factor Loadings for equally-weighed size portfolio with respect to real interest rate ........................................................................................................................ 82
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>RMSE, BIAS and STDC for parameter estimators across ((\sigma^2, \lambda)) with ((n, T) = (100, 5)).</td>
<td>21</td>
</tr>
<tr>
<td>1.2</td>
<td>RMSE, BIAS and STDC for (\hat{\theta}) for ((\sigma^2, \lambda) = (1, 1)) across different sample sizes.</td>
<td>22</td>
</tr>
<tr>
<td>1.3</td>
<td>Effects of Variance Ratio ((\lambda)) on the BIAS, STDC and RMSE of (\hat{\theta}) with ((n, T) = (100, 5)).</td>
<td>23</td>
</tr>
<tr>
<td>1.4</td>
<td>Effects of Sample Size on Smooth Coefficients.</td>
<td>24</td>
</tr>
<tr>
<td>1.5</td>
<td>Empirical relative rejection frequency of the (\hat{T}_n) statistic with Bootstrap.</td>
<td>25</td>
</tr>
<tr>
<td>1.6</td>
<td>List of Countries According to Human Capital (as of 2011)</td>
<td>26</td>
</tr>
<tr>
<td>1.7</td>
<td>Summary Statistics</td>
<td>27</td>
</tr>
<tr>
<td>1.8</td>
<td>Estimation of the Semiparametric Smooth Coefficient Stochastic Frontier</td>
<td>28</td>
</tr>
<tr>
<td>2.1</td>
<td>List of Countries According to Economic Freedom (as of 2010)</td>
<td>64</td>
</tr>
<tr>
<td>2.2</td>
<td>Summary Statistics</td>
<td>65</td>
</tr>
<tr>
<td>2.3</td>
<td>Estimation of the Semiparametric Smooth Coefficient Stochastic Frontier</td>
<td>65</td>
</tr>
<tr>
<td>2.4</td>
<td>Maximum Likelihood Estimation of Parametric Benchmark Models</td>
<td>66</td>
</tr>
<tr>
<td>3.1</td>
<td>Summary Statistics</td>
<td>77</td>
</tr>
<tr>
<td>3.2</td>
<td>Smooth coefficient factor loadings of Fama – French five factor model for equally weighted portfolios</td>
<td>78</td>
</tr>
<tr>
<td>3.3</td>
<td>Smooth coefficient factor loadings of Fama – French five factor model for valued weighted portfolios</td>
<td>79</td>
</tr>
<tr>
<td>3.4</td>
<td>OLS estimate of Fama – French five factor model for equally weighted portfolios</td>
<td>80</td>
</tr>
<tr>
<td>3.5</td>
<td>OLS estimate of Fama – French five factor model for valued weighted portfolios</td>
<td>81</td>
</tr>
</tbody>
</table>
Chapter 1

Semiparametric Smooth Coefficient Stochastic Production Frontier

1.1 Introduction

Since the seminar work of Aigner et al. (1977) and Meeusen and Van Den Broeck (1977), the stochastic frontier (SF) approach as a tool to model and estimate efficiency has grown exponentially (see Greene (1993), Coelli (1995), Kumbhakar and Lovell (2000), Parmeter and Kumbhakar (2014), and Kumbhakar et al. (2015) for extensive reviews of different variants of the SF models and their applications). The SF models are popular among the econometricians due to the fact that these models accommodate stochastic noise (production shocks) and can separate the noise from inefficiency. Furthermore, one can perform statistical tests on many economic hypotheses of interest. In doing so, however, restrictive assumptions are made on either the production frontier function and/or the distributional assumptions on the inefficiency and the noise terms. For example, the SF model pioneered by Aigner et al. (1977) and Meeusen and Van Den Broeck (1977) and extended by many in the last three decades uses a parametric frontier along with a composite error term in which the one-sided inefficiency term follows a particular distribution (half-normal, exponential, truncated normal, gamma, etc.), while the noise term follows a normal distribution. This basic structure is followed in panel data models as well with some extensions on the temporal behavior of inefficiency.

With panel data, one can relax the distribution assumption on the inefficiency term at the expense of some other restrictive assumptions. For example, Schmidt and Sickles (1984) estimate a parametric frontier function by relaxing the distribution assumption in a panel data framework by assuming inefficiency to be time-invariant (parameter). Similarly, Cornwell et al. (1990) and Lee and Schmidt (1993) assume inefficiency to follow a deterministic function of time. Recently Horrace and Parmeter (2011) propose

---

1An alternative to the SF approach is the Data Envelopment Analysis which is comprehensively summarized by Simar and Wilson (2008).
to estimate the inefficiency distribution, while Parmeter et al. (2014) estimate the determinant function of the inefficiency, both without making distributional assumption on the inefficiency term. However, the flexibility in modeling the frontier is still limited to some known parametric functional forms such as the Cobb-Douglas and translog. Even with correctly specified distribution for the composite errors, incorrectly specified frontier can still lead to misleading conclusion regarding the inefficiency levels.

Maintaining the distribution structure in Aigner et al. (1977), Fan et al. (1996) first introduce a nonparametric frontier model and examine properties of the estimator in their simulations. Martins-Filho and Yao (2015) investigate the asymptotic properties of the estimators in Fan et al. (1996) and propose a profile-likelihood based estimator for the nonparametric frontier function where the parameters of the composite error distribution carry no asymptotic bias and are efficient in a class of semiparametric estimators defined in Severini and Wong (1992). On the other hand, Kumbhakar et al. (2007) take a very different approach to model all parameters of the distribution of the composite error as smooth functions of the inputs. They estimate the nonparametric frontier and the distribution parameters with a local likelihood method. Their method is generalized to include discrete regressors in Park et al. (2015).

A common feature of the above methods is that the frontier is fully nonparametric, although in general, the rate of convergence of the proposed frontier estimator is rather slow especially when the number of inputs (conditioning variables) is large. The problem is more severe in Kumbhakar et al. (2007) since all “parameters” are local. It is the well-known curse of dimensionality problem afflicting multivariate kernel based nonparametric estimation. Since it is common to have a large number of variables in frontier models, the accuracy of the asymptotic approximation can be rather poor. In this paper, we propose a smooth coefficient model for the frontier function. It allows a more flexible functional form for the frontier function instead of a linear or a semiparametric partially linear form. The sample size required for estimation is not as demanding as a fully nonparametric frontier model, and therefore likely to be useful to the applied researchers. The smooth coefficient regression model has been popularized by Li et al. (2002) and Cai and Wang (2008). The semiparametric frontier model proposed in this article is different from the standard smooth coefficient regression model because the conditional mean of the composite error is not zero due to the presence of the one-sided inefficiency term. We propose a multi-step estimation procedure to consistently estimate the semiparametric frontier function.

For its general applicability, we consider a panel data framework, assuming that the composite error term follows a certain distribution which depends on some inputs or environment variables $z_{it}$. Unlike the approaches considered in Cornwell et al. (1990), Battese and Coelli (1992), Lee and Schmidt (1993), and Kumbhakar and Wang (2005) which allow the inefficiency to differ across years through a time trend, we allow the inefficiency to vary across individual and time through its dependence on $z_{it}$. We note that it
is important to determine the level of inefficiency and also to understand how the inefficiency is affected by some observed environmental variables. Ignoring the effect of such variables in the composite error term, especially in the one-sided inefficiency term, can cause biased estimates of the frontier function and technical inefficiency level. To capture the dependence of the inefficiency term on $z_{it}$, we allow the conditional mean and variance of the inefficiency term to be a function of $z_{it}$, known up to certain parameters.\footnote{See section 5 in Parmeter and Kumbhakar (2014) for discussion on different methods to account for the determinant of inefficiency.}

We propose a multi-step estimator for the parameters of the composite error distribution and the smooth coefficients in the production frontier function. Under some assumptions, we establish the consistency and asymptotic normality of the proposed estimators to facilitate inference on them. The estimated parameters converge at a rate of $\sqrt{n}$ – the same convergence rate as the parametric estimators based on correct parametric specification of the frontier function. However, the asymptotic bias and variance are impacted by the estimation of the smooth coefficients. On the other hand, the estimator of the smooth coefficients converges at a slower rate which depends only on the dimension of $z_{it}$ and not on that of all the inputs. Furthermore, we propose estimators for the frontier function whose asymptotic distribution is the same as that of the estimators assuming the parameters of the composite error distribution to be known. We illustrate their finite sample performance through a Monte-Carlo study.

The rest of this chapter is organized as follows. In Section 2, we present our multi-step semiparametric estimation for the smooth coefficient frontier function and the parameters in the composite error distribution. Section 3 details their asymptotic characterization. We illustrate their finite sample performance in a Monte-Carlo study in Section 4. The empirical example based on a dataset of 134 countries from 1990 to 2011 is presented in Section 5. Section 6 concludes the paper. Proof of all the theorems, including three supporting lemmas, are provided in the Appendix.

### 1.2 A multi-step semiparametric estimation

We consider the following semiparametric smooth coefficient stochastic frontier model with panel data

$$ y_{it} = \alpha(z_{it}) + X_{it}'\beta(z_{it}) + \epsilon_{it}, \quad i = 1, \ldots, n, t = 1, \ldots, T, $$

where $X_{it} \in \mathbb{R}^p$ are random regressors representing, for example, the traditional input variables, and $z_{it} \in \mathbb{R}^q$ denote the exogenous environmental variables such as R&D, human capital, or even time index. Note that $\alpha(\cdot), \beta(\cdot) = (\beta_1(\cdot), \ldots, \beta_p(\cdot))'$ are unknown smooth functions of environmental variables $z_{it}$ which allows the frontier function to shift in a completely flexible manner. Thus, different technologies...
We follow Wang and Ho (2010) and assume that the composite error term $\epsilon_{it}$ consists of a two-sided error term representing random noise and a one-sided random term representing inefficiency. We also allow the $z_{it}$ variables to affect the distribution of $\epsilon_{it}$. We follow Wang and Ho (2010) and assume that $\epsilon_{it}$ is independent of $X_{it}$ but can depend on $z_{it}$. We adopt this assumption for illustrative purpose, although we conjecture that asymptotic properties of $\alpha(\cdot)$ and $\beta(\cdot)$ estimators will not be altered even when we allow dependence of $\epsilon_{it}$ on both $X_{it}$ and $z_{it}$. For $z_{i} = (z_{i1}, \cdots, z_{iT})'$ and $\epsilon_{i} = (\epsilon_{i1}, \cdots, \epsilon_{iT})'$, we denote the conditional density of $\epsilon_{i}$ given $z_{i}$ by $h(\epsilon_{i}; z_{i}, \theta_{0})$, the conditional mean and variances by $E(\epsilon_{it}|z_{it}) = -\mu(z_{it}; \theta_{0})$ and $V(\epsilon_{it}|z_{it}) = \sigma^{2}(z_{it}; \theta_{0})$, respectively. Since we specify $h(\epsilon_{i}; z_{i}, \theta_{0})$ both $\mu(z_{it}; \theta_{0})$ and $\sigma^{2}(z_{it}; \theta_{0})$ are known functions of $z_{it}$ up to the unknown parameter vector $\theta_{0} \in \mathbb{R}^{d}$. This set-up includes previous efforts to model (i) the conditional variance of the inefficiency term as a function of $z_{it}$ (Caudill and Ford (1993), Caudill et al. (1995), and Hadri (1999)), (ii) the conditional mean (Kumbhakar et al. (1991), Reifschneider and Stevenson (1991), Huang and Liu (1994), and Battese and Coelli (1995)), (iii) a combination of both (Wang (2002)), and (iv) the “scaling” function approach where inefficiency term is modeled as a product of a one-sided random variable and a positive function of $z_{it}$ (Wang and Schmidt (2002)).

However, since $\mu(z_{it}; \theta_{0}) \neq 0$ we can not apply the standard smooth coefficient regression model directly. Instead, given that

$$y_{it} - E(y|z_{it}) = (X_{it} - E(X|z_{it}))'\beta(z_{it}) + \mu(z_{it}; \theta_{0}) + \epsilon_{it}, \tag{1.2}$$

we estimate $\beta(\cdot)$, $\theta_{0}$, and $\alpha(\cdot)$ with the following steps:

1. We estimate $E(y|z)$ by $\hat{\pi}_{y}(z) = \hat{c}_{0}$, $E(X_{j}|z)$ by $\pi_{X_{j}}(z) = \hat{d}_{0}$, using the local linear estimators, viz.,

$$\begin{align*}
(\hat{c}_{0}, \hat{c}_{1}) &= \arg\min_{c_{0}, c_{1}} \sum_{i=1}^{n} \sum_{t=1}^{T} (y_{it} - c_{0} - (z_{it} - z)'c_{1} \right)^2 K\left(\frac{z_{it} - z}{h}\right), \[3.5em]
(\hat{d}_{0}, \hat{d}_{1}) &= \arg\min_{d_{0}, d_{1}} \sum_{i=1}^{n} \sum_{t=1}^{T} (X_{j, it} - d_{0} - (z_{it} - z)'d_{1} \right)^2 K\left(\frac{z_{it} - z}{h}\right).
\end{align*}$$

Here $K(\cdot) : \mathbb{R}^{2} \to \mathbb{R}$ is the kernel function and $h \to 0$ as $n \to \infty$ is the bandwidth.\footnote{We can use different bandwidths and kernel functions in the local linear regressions above for estimating $E(y|z)$ and $E(X_{j}|z)$, but we keep them to be the same for ease of notations.}

Let $\hat{\epsilon}_{it} = \epsilon_{it} + \mu(z_{it}; \theta_{0})$, then $E(\hat{\epsilon}_{i}|z_{it}) = 0$. From Equation (1.2), $y_{it} - E(y|z_{it}) = (X_{it} - E(X|z_{it}))'\beta(z_{it}) + \hat{\epsilon}_{it}$, we construct $\hat{y}_{it} = y_{it} - \pi_{y}(z_{it})$, $\hat{X}_{j, it} = X_{j, it} - \pi_{X_{j}}(z_{it})$ and obtain $\hat{\beta}(z) = (\pi_{y}(z_{it}) - \hat{y}_{it})'(\hat{X}_{j, it})^{-1}$.}

Chapter 1. Semiparametric Smooth Coefficient Stochastic Production Frontier

4
$(\hat{\beta}_1(z), \ldots, \hat{\beta}_p(z))' = b_0$, where $b_0 = (b_{01}, \ldots, b_{0p})'$, $\hat{b}_1 = (\hat{b}_{11}, \ldots, \hat{b}_{1p})'$, $\hat{b} = (\hat{b}_0, \hat{b}_1)'$ such that

$$\hat{b} = \arg \min_b n \sum_{i=1}^n \sum_{t=1}^T (\hat{y}_{it} - \hat{W}_it b)^2 K \left( \frac{z_{it} - z}{h} \right),$$

(1.3)

where $\hat{W}_it = \begin{bmatrix} \hat{X}_it \\ \hat{X}_it \otimes (z_{it} - z) \end{bmatrix}$ is a $[p \times (q + 1)] \times 1$ vector, $\otimes$ denotes the Kronecker product and $\hat{X}_it = (\hat{X}_{i1, it}, \ldots, \hat{X}_{ip, it})'$. 

2. Since $\mu(z_{it}; \theta_0)$ is known up to the parameter vector $\theta_0$, we can construct $\hat{\varepsilon}_{it}(\theta) = \hat{y}_{it} - \hat{\bar{X}}_it \hat{\beta}(z_{it}) - \mu(z_{it}; \theta)$, $\hat{\varepsilon}_i(\theta) = \{\hat{\varepsilon}_{it}(\theta)\}_{t=1}^T$. We then estimate $\theta$ by $\hat{\theta}$ via pseudo-likelihood estimation method with the constructed observations

$$\hat{\theta} = \arg \max_{\theta} \ln \prod_{i=1}^n h(\hat{\varepsilon}_i(\theta); z_i, \theta).$$

(1.4)

3. With $\hat{\theta}$ obtained from above, we estimate $\alpha(z)$ by $\hat{\alpha}(z) = \pi_y(z; h_1) - \sum_{j=1}^p \pi_{X_j}(z; h_1) \hat{\beta}_j(z; h_1) + \mu(z; \hat{\theta})$, where $\pi_y(z; h_1), \pi_{X_j}(z; h_1)$ and $\hat{\beta}_j(z; h_1)$ are defined similar to $\pi_y(z), \pi_{X_j}(z)$ and $\hat{\beta}_j(z)$, except that we use a different bandwidth $h_1$. 

4. Given $\hat{\theta}$ from Step 2, we can estimate $\mu(z_{it}; \hat{\theta})$. For $\delta(z) = (\alpha(z), \beta_1(z), \ldots, \beta_p(z))'$, we observe that $y_{it} + \mu(z_{it}; \theta_0) = (1, X_{it}') \delta(z_{it}) + \hat{\varepsilon}_{it}$ and $E(\hat{\varepsilon}_{it}|X_{it}, z_{it}) = 0$. Motivated by this observation, we construct a new estimator $\hat{\delta}(z) = (\hat{\alpha}(z), \hat{\beta}_1(z), \ldots, \hat{\beta}_p(z))' = \hat{a}_0$, where $\hat{a} = \arg \min_a \sum_{i=1}^n \sum_{t=1}^T (\hat{y}_{it} - \hat{Q}_{it} a)^2 K \left( \frac{z_{it} - z}{h_1} \right)$, $\hat{a} = (\hat{a}_0, \hat{a}_1)'$, $\hat{a}_0 = (\hat{a}_{01}, \ldots, \hat{a}_{0p})'$, $\hat{a}_1 = (\hat{a}_{11}, \ldots, \hat{a}_{1p+1})'$, $\hat{a}_j \in \mathbb{R}^q$, $\hat{y}_{it} = y_{it} + \mu(z_{it}; \hat{\theta})$, and $\hat{Q}_{it} = \begin{bmatrix} (1, X_{it}')' \\ (1, X_{it}')' \otimes (z_{it} - z) \end{bmatrix}$.

Unlike Fan et al. (1996) and Martins-Filho and Yao (2015), we model the production frontier as a semiparametric smooth coefficient function. Furthermore, using panel data, we allow the efficiency score to be dependent on the $z_{it}$ variables. Our estimator $\hat{\beta}(z)$ involves estimates of $E(y|z)$ and $E(X_j|z)$, which have not been dealt with in Cai and Li (2008). On the other hand, $\hat{\beta}(z)$ does not require the treatment of endogenous variables, since $E(\hat{\varepsilon}_{it}|X_{it}, z_{it}) = 0$ in equation (1.2). We take a different approach to construct our asymptotic property for $\hat{\beta}(\cdot)$ by examining each individual component $\hat{\beta}_l(\cdot)$, for $l = 1, \ldots, p$. This approach enables us to have uniform convergence results on $\hat{\beta}_l(\cdot)$ and facilitates our analysis on $\hat{\theta}$ and $\hat{\alpha}$.

Of course, we provide a multivariate asymptotic statement of $\hat{\beta}(\cdot)$ in Corollary 1 as well.

We use two different bandwidths in constructing $\hat{\alpha}(\cdot)$ which depends on the estimates $\hat{\theta}$. Careful choices of bandwidths in Assumption A6(2), where we use a slightly under smoothed bandwidth $h$, make sure that the bias introduced in estimating $\theta$ does not carry through in the asymptotic distribution of $\hat{\alpha}(\cdot)$ as in Theorem 3. However, our asymptotic analysis (Theorem 3) shows that many additional
characteristics of the DGP appear in the bias of \( \hat{\alpha}(-) \), which motivates us to consider a much simplified estimator \( \hat{\alpha}(-) \) in step 4. This simplification results in a much simpler bias expression (see Theorem 4).

Our set-up on the distribution of the composite error allows the distribution of the one-sided error to be the half-normal, truncated normal, exponential or gamma. For its popularity, we consider \( v_{it} \sim IIDN(0, \sigma_v^2) \), and for a known nonnegative function \( g(z_{it}; \eta) \) up to unknown parameters \( \eta \), \( u_{it} = u_i g(z_{it}; \eta) \), and \( u_i \sim IID[N(0, \sigma_u^2)] \) independent of \( z_{it} \). Thus, the one-sided inefficiency term exhibits the “scaling” property (Wang and Schmidt (2002)), where the shape of \( u_{it} \) is fixed at \( u_i \) (which captures the base inefficiency level) and the scale of \( u_{it} \) is controlled by \( g(z_{it}; \eta) \) (which captures the impacts of the environment variables \( z_{it} \)). Furthermore, when \( g(z_{it}; \eta) \) is the exponential function as we specify below in the simulation, \( \eta \) can be interpreted as the semi-elasticity of expected inefficiency with respect to \( z_{it} \), no matter what distribution assumption is placed on \( u_i \). Then we have \( \mu(z_{it}; \theta_0) = \sqrt{\pi} \sigma_u g(z_{it}; \eta) \) and \( \sigma^2(z_{it}; \theta_0) = \sigma^2_u + \pi^2 \sigma^2_u \theta^2(z_{it}; \eta) \). Let \( \sigma^2_{\eta} = \sigma^2_u + \pi^2 \sum_{t=1}^T g^2(z_{it}; \eta) \), \( \lambda = \sigma_u / \sigma_v \), \( \mu_{\eta} = -\sqrt{\pi} \sum_{t=1}^T \epsilon_{it} g(z_{it}; \eta) / \sigma^2_{\eta} \), \( \sigma^2_{\eta} = \sigma^2_u \sigma^2_{\eta} / \sigma^2_u \), then \( \sigma^2_{\eta} = \sigma^2_{\eta} / (1 + \lambda^2 \sum_{t=1}^T g^2(z_{it}; \eta)) \). Denoting the PDF and CDF of a standard normal by \( \phi(-) \) and \( \Phi(-) \), we follow Pitt and Lee (1981) to obtain

\[
h(\epsilon_i; z_i, \theta) = \frac{2}{\sigma_v} [1 - \Phi(- \frac{\mu_{\eta}}{\sigma_v})] \prod_{t=1}^T \phi(\frac{\epsilon_{it}}{\sigma_v}) \exp\left(\frac{1}{2} \frac{\mu_{\eta}^2}{\sigma_v^2}\right).
\]

For \( \theta = (\sigma^2_u, \sigma^2_v, \eta)' \) and for some constant \( C \) that does not depend on \( \theta \), we obtain

\[
\ln \prod_{i=1}^n h(\epsilon_i; z_i, \theta) = C - \frac{n(T-1)}{2} \ln \sigma_v^2 - \frac{n}{2} \sum_{i=1}^n \ln(\sigma^2_v + \sigma^2_u \theta^2(z_{it}; \eta)) + \sum_{i=1}^n \ln[1 - \Phi(- \frac{\mu_{\eta}}{\sigma_v})] + \frac{n}{2} \sum_{i=1}^n \frac{\mu_{\eta}^2}{\sigma_v^2} - \frac{n}{2} \sum_{i=1}^n \epsilon_{it}^2 (\theta).
\]

### 1.3 Asymptotic Characterization

We provide the asymptotic properties of our proposed semiparametric estimators with the following assumptions. Below we denote a generic constant by \( C \), the magnitude of which is inconsequential for the asymptotic analysis and can vary from one place to another. Let’s denote a generic function \( g(z) \) and all of its partial derivatives of order \( \leq 2 \) are continuous and uniformly bounded on \( \mathbb{R}^q \).

**Assumptions:**

\( \text{A1:} \{ Z_{it}, X_{it}, \epsilon_{it} \} \) are identically and independently distributed (IID) across \( i \) for each fixed \( t \) and strictly stationary over \( t \) for each fixed \( i \).

**A2:** \( K(z) : \mathbb{R}^q \rightarrow \mathbb{R} \) is a product kernel \( K(z) = \prod_{j=1}^q K(z_j) \) with symmetric \( K(z) : \mathbb{R} \rightarrow \mathbb{R} \) such that: (1) \( |K(z)|z^j| \leq C \) for all \( z \in \mathbb{R} \) and \( j = 0, 1, 2, 3 \); (2) \( \int |z^j K(z)|dz \leq C \) for \( j = 0, 1, 2, 3 \); (3) \( \int K(z)dz = 1 \), \( \int zK(z)dz = 0 \), \( \int z^2 K(z)dz = \mu_K < \infty \); (4) \( K(z) \) is continuously differentiable on \( \mathbb{R} \) with \( |z^j \frac{d}{dz} K(z)| \leq C \) for all \( z \in \mathbb{R} \) and \( j = 0, 1, 2, 3 \).
A3: (1) Denote the marginal density of $z$ by $f_z(z)$ and let $f_z(z) \in C^2$; (2) $0 < \inf_{z \in \mathbb{R}^q} f_z(z)$, where $\mathcal{G}$ is a compact subset of $\mathbb{R}^q$. (3) Denote the joint density of $z_1, z_t$ by $f_{z_1, z_t}(z_1, z_t)$, and it is continuous over $\mathbb{R}^q \times \mathbb{R}^q$. (4) Denote the conditional density of $z$ given $X$ by $f_{z \mid X}(z)$, and $f_{z \mid X}(z) < \infty$. (5) Denote the conditional density of $z$ given $\epsilon$ by $f_{z \mid \epsilon}(z)$, and $f_{z \mid \epsilon}(z) < \infty$. (6) Denote the conditional density of $z$ given $X$ and $\epsilon$ by $f_{z \mid X, \epsilon}(z)$, and $f_{z \mid X, \epsilon}(z) < \infty$.

A4: (1) $E(\epsilon_t \mid X_t, z_t) = E(\epsilon_t \mid X_t, z_t) + \mu(z_t; \theta_0) = 0$. (2) $E(\epsilon_t^2 \mid z_t) = E(\epsilon_t^2 \mid z_t) < C$ for some $\delta > 0$. (3) $E(\epsilon_t^2 \mid z_t)$ is continuous at $z_t$. (4) $E((X_{l,t}^\theta)^j (X_{l,t+1}^\theta)^{j'} (\epsilon_t \epsilon_{t+1})^{j''} | z_t_1, z_{t+1}) < C$ and continuous at $z_t, z_{t+1}$ for $j, j' = (0, 1), j'' = (0, 1), t = 1, \ldots, T - 1$ and $l, l' = (1, \ldots, p)$. (5) $\forall l = 1, \ldots, p, E(|X_{l,t}^\theta |^a) < \infty$ for some $a > 2$, where $X_{l,t}^\theta = X_t - E(X_t | z_t)$. (6) $E(|X_t^2 |^a | z_t) < C$ for some $\delta > 0$. (7) $E(X_{l,t}^2 | z_t)$ is continuous at $z_t$. (8) $v_{W}(z) = E(X_{l,t}^\theta X_{l,t}^\theta | z_t) \forall l, l' = 1, \ldots, p$, and $v_{W}(z) \in C^2$.

$\Omega(z) = \begin{pmatrix} 1 & \frac{E(X_t^2 | z_t)}{E(X_t) | z_t} \end{pmatrix}$ is a positive definite matrix.

A5: (1) $E(X_t | z_t) \in C^2$. (2) $\alpha(z) \in C^2$. (3) $\beta_t(z) \in C^2$. (4) $\mu(z; \theta_0) = -E(\epsilon(z, X)) = -E(\epsilon(z)) \in C^2$.

A6: (1) $h = Cn^{-1/(q+4)}$ for $q < 8$, $n \to \infty$ and $T$ is fixed. (2) $h = Cn^{-1/(q+4)} - \delta$ for $q < 8$ and some $\delta > 0$. (3) $h = Cn^{-1/(q+4)} - \delta$, $n \to \infty$ and $T$ is fixed.

B1: (1) $\theta_0 \in int(\Theta)$, the interior of the compact set $\Theta \subset \mathbb{R}^d$. (2) $\forall \theta \in \Theta$, if $\theta \neq \theta_0$, then $h(\epsilon_t; z_t, \theta) = h(\epsilon_t; z_t, \theta_0)$. (3) If $\{\theta_i\}_{i=1, \ldots}$ is a sequence in $\Theta$ s.t. $\theta_i \to \theta$ as $i \to \infty$, then $h(\epsilon_t; \theta_i) \to h(\epsilon_t; \theta)$ as $\forall \theta \in \Theta$. (4) $h(\sup_{\theta \in \Theta} | h(\epsilon_t; z_t, \theta) |) < \infty$.

B2: $\ln h(\epsilon_t(z_t, \theta); z_t, \theta) - \ln h(\epsilon_t(z_t); z_t, \theta) | \leq (1 + o_p(1))b_0(y_t, X_t, z_t, \theta) \{ \sum_{i=1}^{p} b_i(y_t, X_t, z_t, \theta) \} \leq E(\theta_t | x_t) + \sum_{j=1}^{p} \beta_j(z_t) - \beta_j(z_t) | b_j(y_t, X_t, z_t, \theta) |, \text{ where } l = 1, 2 \text{ and } \tau = 0, 1, \text{ E}(\sup_{\theta \in \Theta} | b_0(y_t, X_t, z_t, \theta) \sum_{i=1}^{p} b_i(y_t, X_t, z_t, \theta) |) < C.

B3: (1) $h(\epsilon; z, \theta)$ is twice continuously differentiable with respect to (w.r.t.) $\theta$ and $h(\epsilon; z, \theta) > 0$ on some open ball $S_0, \theta = S(\theta_0, d(\theta_0)) \subset \Theta$. (2) $h(\epsilon_t; z_t, \theta)$ is twice continuously differentiable w.r.t. $\epsilon_t$. For fixed $z_t, \theta_0 = \alpha(z_t) + E(X_t | z_t)^\theta \beta(z_t), \text{ and } \epsilon_t = y_t - (X_t^\theta)^\theta \beta(z_t) - \eta_0$. For $H$, a compact subset of $\mathcal{R}$, let $S_{\eta, \theta} = S(\eta_0, d(\eta_0))$, an open interval in $H$ containing $\eta_0$. Assume $\sup_{\theta \in \Theta} \int \ln h(\epsilon_t(z_t, \theta)) < C$ and $\sup_{\theta \in \Theta} \sum_{i=1}^{p} b_i(y_t, X_t, z_t, \theta) < C$. (3) $\mu(z_t; \theta)$ is twice continuously differentiable in $\theta$, $\mu(z_t; \theta) < C$. (4) Assume $\frac{\partial}{\partial \theta} \ln h(\epsilon; z, \theta)$ is continuously differentiable in $\theta$, for $z_{m, t} = \{z_{m, t}^j \mid j = 0, 1, \ldots, m\}$, $E(\sup_{\theta \in S_{\eta, \theta}} | \frac{\partial^2}{\partial \theta \partial \theta} \ln h(\epsilon; z, \theta) | E(1 + | f(\epsilon_t + \mu(z_t; \theta) h(\epsilon_t; z_{m, t}; \theta) d\epsilon_t) |) < C.$
\( E( \sup_{\theta \in S_{0, \theta}} | \frac{\partial}{\partial \theta} \mu(z_{it}; \theta) | E[ \int \frac{\partial}{\partial \theta} \ln h(\epsilon; z_{i}, \theta) h(\epsilon; z_{i}, \theta_{0}) d\epsilon_i ] ) < C, \) 
\( \int \sup_{\theta \in S_{0, \theta}} | \frac{\partial}{\partial \theta} \ln h(\epsilon; z_{i}, \theta) \ln h(\epsilon; z_{i}, \theta_{0}) d\epsilon_i < C. \)

(5) Let \( Q_{it} = \frac{\partial}{\partial \theta} \ln h(\epsilon_{i}; z_{i}, \theta_{0}) + \frac{\tau}{i \in T} \frac{\partial}{\partial \theta} \ln h(\epsilon_{i}; z_{i}, \theta_{0}) \frac{\partial}{\partial \theta} \mu(z_{i}; \theta_{0}) + \frac{\tau}{i \in T} (-\epsilon_{i}) \times \int \int \frac{\partial^{2}}{\partial x_{m_{i}}^{2}} \ln h(\epsilon_{m_{i}}; \epsilon_{m_{i}}, \theta_{0}) h(\epsilon_{m_{i}}; \epsilon_{m_{i}}, \theta_{0}) d\epsilon_{m_{i}} \frac{f_{km_{i}}(z_{m_{i}, \epsilon_{i}})}{f_{k}(z_{i})} d\epsilon_{m_{i}} \cdot d\epsilon_{m_{i}, T}, \)
we assume \( \sigma_{p}^{2} = E(Q_{it}Q_{it}^{T}) \) exists and is positive definite and we denote the joint density of \( z_{m} \) by \( f_{zm}(\cdot) \).

(6) Let \( \bar{H} = E(\frac{\partial^{2}}{\partial \theta^{2}} \ln h(\epsilon_{i}; z_{i}, \theta_{0})) - TE(\frac{\partial}{\partial \theta} \mu(z_{i}; \theta_{0}) \frac{\partial^{2}}{\partial x_{m_{i}}^{2}} \ln h(\epsilon_{i}; z_{i}, \theta_{0}) + \frac{\partial^{2}}{\partial \theta \partial \theta} \ln h(\epsilon_{i}; z_{i}, \theta_{0}) \frac{\partial}{\partial \theta} \mu(z_{i}; \theta_{0})). \)

We assume that \( \bar{H} \) exists and is nonsingular.

**B4:** (1) \( E( \sup_{\theta \in S_{0, \theta}} | \frac{\partial^{2}}{\partial \theta^{2}} \ln h(\epsilon_{i}; z_{i}, \theta_{0}) (C + \mu(2)(z_{i}; \theta)) | ) ) < C \) for \( a, b = 1, \cdots, d. \) (2) \( \frac{\partial^{2}}{\partial \theta \partial \theta} \ln h(\epsilon_{i}; z_{i}, \theta_{0}) \) is continuously differentiable on \( S_{0, \theta} \). (3) We assume that \( E(\frac{\partial^{2}}{\partial \theta \partial \theta} \ln h(\epsilon_{i}; z_{i}, \theta_{0})) \) for \( \tau = 0, 1, E(\frac{\partial^{2}}{\partial \theta \partial \theta} \ln h(\epsilon_{i}; z_{i}, \theta_{0}) | \theta) \) \) \( E(\frac{\partial}{\partial \theta} \ln h(\epsilon_{i}; z_{i}, \theta_{0}) | \theta) \), \( E(\frac{\partial^{2}}{\partial \theta \partial \theta} \ln h(\epsilon_{i}; z_{i}, \theta_{0}) | \theta), \) \( E(\frac{\partial^{3}}{\partial \theta \partial \theta} \ln h(\epsilon_{i}; z_{i}, \theta_{0}) | \theta) \) are continuous at \( \theta \in S_{0, \theta} \).

Assumption A1 requires that the observations are IID across \( i \) and stationary across \( t \), a standard assumption in the panel data literature. A2 gives some standard moment and smoothness conditions on the kernel function. We focus attention on the popular second order kernel, so the Gaussian kernel can be used. Correspondingly in A5, we assume that the functions to be estimated are smooth up to the second degree, so we do not explore the potential gain in relaxing the assumptions from using higher order kernel and requiring higher degree of the smoothness of the functions. A3 asks that the marginal and conditional density of \( z \) are smooth and bounded, which allows us to perform asymptotic analysis of estimators of \( \alpha(\cdot) \) and \( \beta(\cdot) \) using the result in Lemma 2 repeatedly. A4 (1)-(4) places conditional moment conditions on \( \epsilon_{it}, (5)-(8) \) list boundedness and smooth condition on the conditional moment involving \( X, \) and (9)-(10) allow identification of the \( \alpha(\cdot) \) and \( \beta(\cdot) \) in the frontier. Finally, A6 allows that \( n \) goes to infinity while \( T \) is fixed. Under this set-up, we conjecture that our asymptotic analysis continue to hold even without the stationarity assumption in A1, but we keep the current assumption to be in-line with previous analysis and the conclusion will be concise. Though the asymptotic analysis can be carried out with more general requirement on the bandwidths, we focus attention on the ones listed in A6, so specifically, the optimal bandwidths for estimating \( \alpha(\cdot) \) and \( \beta(\cdot) \) in a smooth coefficient regression model can be used in the estimation of the frontier functions.

Assumptions B1-B4 allow investigation of the parameter estimators in the composite error distribution. B1 guarantees that a unique maximum of \( E(\ln h(\epsilon; z, \theta)) \) exists at \( \theta_{0} \) (see Theorem 2.5 in Newey and McFadden (1994)). To obtain consistency of \( \hat{\theta} \) which depends on \( \hat{\beta}(\cdot) \), we further need B2 to control the magnitude of the error generated in the first step nonparametric estimation. With Assumptions A4,
A5 and the fact that $\phi(x)/\Phi(x)$ is convex, goes asymptotically to $-x$ as $x \to -\infty$, and goes to zero as $x \to \infty$, we can show that $B2$ is satisfied by the normal and half-normal composite error density considered in last section. To obtain the asymptotic normality of $\hat{\theta}$, we assume B3 and B4 that higher order derivatives of $\ln h(\epsilon; z, \theta)$ are smooth and bounded.

Theorem 1 provides the asymptotic normality result for the first step estimator $\hat{\beta}(\cdot)$.

**Theorem 1.** With Assumptions A1-A5, A6(1), for $V_{pp}(z)$ defined in Assumption A4, $V_{ip}(z)$ being the $(p-1) \times (p-1)$ matrix obtained by deleting the $l$th row and column from $V_{pp}(z)$, and $|V_{ip}(z)|$ denotes its determinant, we have for $l = 1, \cdots, p$,

$$
\sqrt{n}Th\hat{\beta}_l(z) - \beta_l(z) \sim N\left(0, \int K(\psi)\psi^T \beta_l^{(2)}(z) \psi d\psi + o_p(h^2)\right).
$$

Now let’s consider the joint asymptotic distribution of $\hat{\beta}(z) - \beta(z)$ below. Note the asymptotic distribution is similar to Theorem 4 of Cai and Li (2008). So the additional steps of estimating $E(y|z)$ and $E(X_j|z)$ in $\hat{\beta}(z)$ does not impact the asymptotic distribution.

**Corollary 1.** With Assumptions A1-A5, A6(1), we have

$$
\sqrt{n}Th\hat{\beta}(z) - \beta(z) = \frac{h^2}{2} \int K(\psi)\psi^T \beta^{(2)}(z) \psi d\psi + o_p(h^2) \Rightarrow N(0, \int K^2(\psi)\psi d\psi \frac{V(\epsilon|z)V_{ip}(z)}{F(z)V_{pp}(z)}).
$$

Now we present the asymptotic normality of the parameter estimator $\hat{\theta}$, which is obtained at a parametric rate of $\sqrt{n}$. Note that the first step estimation does have an impact on its asymptotic bias as well as its asymptotic variance.

**Theorem 2.** Denote the true value of $\theta$ by $\theta_0$, then

(I) Under Assumptions A1-A5, A6(1), B1-B2, we have $\hat{\theta} \overset{d}{\to} \theta_0$.

(II) Under Assumptions A1-A5, A6(1), B1-B4, for $\bar{H}$ and $\sigma^2_z$ defined in Assumption B3, define

$$
BQ_{2n1} = \frac{h^2}{2} T \int \int \frac{\partial^2}{\partial \theta \partial \epsilon^2} \ln h(\epsilon_i; z_i, \theta_0) h(\epsilon_i; z_i, \theta_0) d\epsilon_i K(\psi) \psi' \times \left\{ \sum_{j=1}^p (-m_{X_j}(z_i) + E^{(2)}(X_j|z_i)\beta_j(z_i)) - \alpha^{(2)}(z_i) + \mu^{(2)}(z_i; \theta_0) \psi f_z(z_i) dz_i d\psi, \right. $$

we have $\sqrt{n}(\hat{\theta} - \theta_0 + \bar{H}^{-1}BQ_{2n1} + o_p(h^2)1_d) \overset{d}{\to} N(0, \bar{H}^{-1}\sigma^2_z \bar{H}^{-1})$.

The asymptotic property of $\hat{\alpha}$ is presented in Theorem 3 below.

**Theorem 3.** With Assumptions A1-A5, A6(2), B1-B4, let

$$
W_{it} = \bar{\epsilon}_i (1 - \sum_{j=1}^p \frac{E(X_j|z)|V_{pp}(z)|}{|V_{ip}(z)|}(X_{ji} - V_{ji} V_{pj}^{-1} S_{i|j})),
$$

Note that the first step estimation does have an impact on its asymptotic bias as well as its asymptotic variance.
where $S^o_{itj} = (\tilde{X}^o_{1,it}, \cdots, \tilde{X}^o_{j-1,it}, \tilde{X}^o_{j+1,it}, \cdots, \tilde{X}^o_{p,it})'$, then we have
$$\sqrt{nTh_1^q}(\hat{\alpha}(z) - \alpha(z) - \frac{h_1^2}{2} \int K(\psi)\psi' \beta_j(z)\beta_j(z)' + \alpha(z) - \mu_2(z;\theta_0)]d\psi + o_p(h_1^2)) \xrightarrow{d} N(0, \int K^2(\psi)d\psi \frac{E(W^2_{f(z)})}{f(z)})$$

The asymptotic bias of $\hat{\alpha}(\cdot)$ depends on many additional terms in the data generating process, beyond the second order derivative of $\alpha(\cdot)$. We propose an estimator $\hat{\alpha}(\cdot)$ and $\hat{\beta}(\cdot)$ in Step 4 for which the asymptotic distribution in Theorem 4 yields a much simplified bias expression in $\hat{\alpha}(\cdot)$.

**Theorem 4.** With Assumptions A1-A5, A6(2), B1-B4, let $\Omega(z) = \begin{bmatrix} 1 & E(X'|z) \\ E(X|z) & E(XX'|z) \end{bmatrix}$, and $Tr(\delta_2(z)) = [Tr(\alpha_2(z)), Tr(\beta_1(z)), \cdots, Tr(\beta_p(z))]'$, where $Tr(A)$ is the trace of matrix $A$, then we have
$$\sqrt{nTh_1^q}[\delta(z) - \delta(z) - \frac{h_1^2}{2} \mu_{K,q}Tr(\delta_2(z)) + o_p(h_1^2)] \xrightarrow{d} N(0, \Omega(z)^{-1} \int K^2(\psi)d\psi \frac{V(z)}{f(z)}).$$

**Comment 1:** from $\Omega(z)^{-1} = \begin{bmatrix} [1 - E(X'|z)](E(XX'|z))^{-1}E(X|z)]^{-1} - \frac{[E(XX'|z)]^{-1}E(X|z)]}{1 - E(X|z)(E(XX'|z))^{-1}E(X|z)} & [V_{pp}(z)]^{-1} \\ \frac{[E(XX'|z)]^{-1}E(X|z)]}{1 - E(X|z)(E(XX'|z))^{-1}E(X|z)} & [V_{pp}(z)]^{-1} \end{bmatrix}$, where $V_{pp}(z) = E((X - E(X|z))(X - E(X|z))'z)$, we obtain the asymptotic variance of $\hat{\alpha}(z)$ as $\int K^2(\psi)d\psi \frac{V(z)}{f(z)}[1 - E(X'|z)](E(XX'|z))^{-1}E(X|z)]^{-1}$, and that of $\hat{\beta}(z)$ as $\int K^2(\psi)d\psi \frac{V(z)}{f(z)}[V_{pp}(z)]^{-1}$.

**Comment 2:** from Theorem 3 and with some algebra, we obtain the asymptotic variance of $\hat{\alpha}(z)$ as $\int K^2(\psi)d\psi \frac{V(z)}{f(z)}[1 - E(X'|z)](E(XX'|z))^{-1}E(X|z)]^{-1}$, thus $\hat{\alpha}(z)$ and $\hat{\alpha}(z)$ possess the same asymptotic variances. If we assume the same bandwidth $h_1$ is used, then the difference between the asymptotic bias of $\hat{\alpha}(z)$ and $\hat{\alpha}(z)$ is $\frac{h_1^2}{2} \mu_{K,q}[Tr(\sum_{j=1}^p 2E(1)(X_j|z)\beta_jz') + \mu_2(z;\theta_0)] + o_p(h_1^2)$, whose sign and magnitude depend on much more detailed characteristic of the data generating processes, which appears only in the bias of $\hat{\alpha}(z)$. On the other hand, the asymptotic variance and bias of $\hat{\beta}(z)$ and $\hat{\beta}(z)$ are identical if we adopt the same bandwidth. Thus we expect potential gains by estimating $\alpha(z)$ with $\hat{\alpha}(z)$ in terms of a much simplified bias expression.

The smooth coefficient function estimators converge at a rate of $\sqrt{nTh_1^q}$, thus suffer from the curse of dimensionality. However, the slower convergence rate depends only on the dimension of $z_{it}$ and not on that of all inputs. Corollary 1 and Theorem 4 further reveal that the asymptotic distributions of $\hat{\beta}(\cdot)$ and $\hat{\beta}(\cdot)$ are the same as those of the estimators assuming the parameters of the composite error distribution to be known.
1.4 Monte Carlo Study

1.4.1 Estimation Procedure

In this section, we perform a Monte Carlo simulation to investigate the finite simple performances of our estimators proposed in Section 2. We have three objectives in performing our simulations. First, we explore whether our estimators \( \hat{\beta} \) are robust across different functional form specifications and different values of parameter specifications. Second, we investigate the effects of various sample sizes (by choosing different combinations of \( n \) and \( T \)) and the effects of different values of \( \lambda (=\sigma_u/\sigma_v) \) on the performance of \( \hat{\beta} \). Third, we examine the effects of different sample sizes on the estimation of smooth coefficients \( \alpha(\cdot) \) and \( \beta(\cdot) \).

We consider the following data generating processes (DGPs) for \( j = 1, 2 \),

\[
\text{DGP } j : y_{it} = \alpha_j(z_{it}) + x_{1,it}\beta_1(z_{it}) + x_{2,it}\beta_2(z_{it}) + v_{it} - u_i g(z_{it}; \eta),
\]

where \( i = 1, 2, \ldots, n \) and \( t = 1, 2, \ldots, T \). In DGP 1, we let \( \alpha_1(z_{it}) = -\frac{1}{1+3z_{it}}, \beta_{11}(z_{it}) = (2z_{it})^3 \) and \( \beta_{21}(z_{it}) = \ln(5z_{it}) \). In DGP 2, \( \alpha_2(z_{it}) = -5\cos(4z_{it}), \beta_{12}(z_{it}) = \sin(4z_{it}) \) and \( \beta_{22}(z_{it}) = \ln\left(\frac{z_{it}}{1.1}\right) \). In both cases, \( g(z_{it}; \eta) = e^{\eta z_{it}} \) where we fix \( \eta \) to be 1, and we generate random variables \( x_{1,it}, x_{2,it} \) and \( z_{it} \) from \( N(1, 0.25^2) \), \( U(1, 2) \) and \( U(0, 1) \), respectively. Finally, \( v_{it} \) and \( u_i \) are drawn from \( N(0, \sigma_v^2) \) and \( |N(0, \sigma_u^2)| \), respectively. To give a clear picture of the smooth coefficient functions used in the simulation, we plot the functions against \( z_{it} \) in Figure 1. We observe that they exhibit substantial nonlinearity over the support of \( z_{it} \), suitable to capture nonlinearities in the production frontiers.

For the estimation, we choose a Gaussian kernel function throughout the simulations and use a simple rule-of-thumb bandwidth of \( h = h_1 = C \hat{\sigma}_{z_{it}} n^{-1/5} \), where \( \hat{\sigma}_{z_{it}} \) is the sample standard deviation of \( z_{it} \) and we let \( C = 1 \). Although our theoretical results in Theorems 3 and 4 call for bandwidths of different magnitudes, we adopt the choices \( h = h_1 \) for simplicity, since in separate simulations using different bandwidths, we observe experiment outcomes of qualitatively similar nature. For each experiment design, we perform 500 repetitions and report the bias (BIAS), standard deviation (STDC) and root mean squared error (RMSE) for \( \hat{\beta} = (\hat{\sigma}_u^2, \hat{\sigma}_v^2, \hat{\eta})' \). We define \( \sigma^2 = \sigma_v^2 + \frac{\pi^2-2}{2}\sigma_u^2E[g^2(z_{it}; \eta)] \) and \( \lambda = \sigma_u/\sigma_v \), where \( E[g^2(z_{it}; \eta = 1)] = \frac{e^\eta-1}{2} \). The true values of \( \sigma_u^2 \) and \( \sigma_v^2 \) are determined once we select \( \sigma^2 \) and \( \lambda \).

To investigate the first objective, we follow Aigner et al. (1977) and Fan et al. (1996) to choose \( (\sigma^2, \lambda) = (1.88, 1.66), (1.63, 1.24) \) and \( (1.35, 0.83) \) as experiment designs for both DGP 1 and 2, and let \( n = 100 \) and \( T = 5 \). Once \( \hat{\beta} \) are obtained, we calculate \( \hat{\lambda} = \hat{\sigma}_u/\hat{\sigma}_v \) and \( \hat{\sigma}^2 = \hat{\sigma}_v^2 + \frac{\pi^2-2}{2}\hat{\sigma}_u^2 \sum_{i=1}^n \sum_{t=1}^T g^2(\hat{\eta}_{z_{it}})/nT \). We report the BIAS, STD and RMSE for \( \hat{\beta}, \hat{\sigma}^2 \) and \( \hat{\lambda} \) in Table 1, in Panel A for DGP 1 and Panel B for DGP 2. The performances are fairly stable across different designs in both DGPs, thereby illustrating that
our parameter estimators based on nonparametric estimation of the smooth coefficients are robust across the three pairs of \((\sigma^2, \lambda)\) and different functional forms. Although our choices of nonlinear production functions are unlike those in Fan et al. (1996), our estimators are generally comparable in magnitudes with theirs. Specifically, when both \((\sigma^2, \lambda)\) decrease, the performances of \(\hat{\sigma}^2, \hat{\lambda}, \hat{\sigma}_u^2\) and \(\hat{\sigma}_v^2\) improve in terms of smaller RMSE, BIAS and STDC – roughly matching the pattern observed in Fan et al. (1996) except for \(\hat{\sigma}_u^2\). However, the performance of \(\hat{\eta}\) improves with larger values of \((\sigma^2, \lambda)\).

The above results are only for \(n = 100\) and \(T = 5\) and \(\lambda\) between 0.83 and 1.66. Now we investigate the performances of our parameter estimators for different sample sizes and for a wide range of \(\lambda\). To explore the effects of different sample sizes on \(\hat{\theta}\), we choose \(n = (50, 100, 200)\) and \(T = (5, 10)\) and follow Olson et al. (1980) and Fan et al. (1996) to set \(\lambda = 1\) and \(\sigma^2 = 1\) for the ease of illustrations. The summarized results of our parameter estimators in Table 2 indicate that the performances of estimators improve in terms of smaller RMSE, BIAS and STDC with larger \(n\) for either \(T = 5\) or 10. The observations are consistent with our Theorem 2, indicating that the parameter estimators are consistent across two different DGP's. We further examine the impact of \(\lambda\) on the performance of \(\hat{\theta}\) for both DGP's in Table 3 for \(n = 100\) and \(T = 5\). Following Fan et al. (1996), we choose \(\lambda = 10^{-1}, 10^{-3/4}, 10^{-2/4}, 10^{-1/4}, 1, 10^{1/4}, 10^{2/4}, 10^{3/4}, 10\) and set \(\sigma^2 = 1\). Although the experiment designs are different, we observe that the performances of \(\hat{\sigma}_u^2\) and \(\hat{\sigma}_v^2\) are comparable with those of Fan et al. (1996) in the sense that although we do not observe a clear pattern in the performances of \(\hat{\sigma}_u^2\) on \(\lambda\), the STDC of \(\hat{\sigma}_v^2\) decreases with larger \(\lambda\). Here, we observe that \(\hat{\eta}\) performs better in terms of smaller RMSE when \(\lambda\) increases in its considered range.

Last, we evaluate the finite sample performances of estimators of our smooth coefficients \(\hat{\alpha}(\cdot), \hat{\beta}_1(\cdot), \hat{\beta}_2(\cdot), \hat{\alpha}(\cdot)\), and \(\hat{\alpha}(\cdot), \hat{\beta}_1(\cdot), \hat{\beta}_2(\cdot)\) at 99 equally spaced grid points on \((0, 1)\), the support of \(z_{it}\). Though we obtain results for all experiment designs as in Tables 1-3, the performances of the smooth coefficient estimates are fairly robust and do not change significantly across different experiment designs. To save space, we report in Table 4 the square root of the average mean squared error (RAMSE) and the average bias (ABIAS) at the grid points over 500 repetitions, using only the experiment designs as in Table 2, i.e., at different sample sizes for both DGP’s by fixing \((\sigma^2, \lambda) = (1, 1)\). It is clear that the performances of all estimators improve with smaller RAMSE, ABIAS when \(n\) increases in both DGP’s for either \(T = 5\) or 10. The average standard deviation for all estimators, which is not reported here, also clearly decreases as \(n\) increases. The observation confirms our Theorems 1, 3 and 4 that \(\hat{\alpha}(\cdot), \hat{\beta}_1(\cdot), \hat{\beta}_2(\cdot), \hat{\alpha}(\cdot)\), and \(\hat{\alpha}(\cdot), \hat{\beta}_1(\cdot), \hat{\beta}_2(\cdot)\) are consistent. We also observe that the bias of \(\hat{\alpha}(\cdot)\) is always smaller than that of \(\hat{\alpha}(\cdot)\). This confirms our observation in Comment 2 that the bias expression of \(\hat{\alpha}(\cdot)\) is much simplified than that of \(\hat{\alpha}(\cdot)\), which can be translated into \(\hat{\alpha}(\cdot)\)'s better finite sample performance in terms of smaller ABIAS. The performances of \(\hat{\alpha}(\cdot), \hat{\beta}_1(\cdot), \hat{\beta}_2(\cdot), \hat{\alpha}(\cdot), \hat{\beta}_1(\cdot), \hat{\beta}_2(\cdot)\) are fairly close to each other in terms of RAMSE where differences
frequently occur in the second or third decimal point, with \( \hat{\alpha}(\cdot), \hat{\beta}_1(\cdot), \hat{\beta}_2(\cdot) \) being slightly preferred at least in our simulation set-up.

Overall, the simulation results indicate our proposed estimators, \( \hat{\theta} \) for the parameters, \( \hat{\alpha}(\cdot), \hat{\beta}(\cdot) \), \( \check{\alpha}(\cdot), \check{\beta}(\cdot) \) for the smooth coefficient production frontier, demonstrate good finite sample performances. Their performances are improving with sample sizes, and are fairly robust across different parameter values of \((\sigma^2, \lambda)\) and various nonlinear functional forms for \( \alpha(\cdot) \) and \( \beta(\cdot) \).

1.4.2 Testing Procedure

In practice, one would like to test whether certain parametric specifications of the production frontier hold. In this section, we propose a modified test for correct parametric specification of \( \delta(z) = (\alpha(z), \beta(z))' \) based on Li et al. (2002) and Li and Racine (2010). We note that besides considering the panel data framework, our set-up differs from theirs in that the conditional mean of \( \epsilon_{it} \) given \( X_{it} \) and \( z_{it} \) is not zero in (1.1) due to the presence of the inefficiency term, and our estimators \( (\hat{\alpha}(\cdot), \hat{\beta}(\cdot)) \) and \( (\check{\alpha}(\cdot), \check{\beta}(\cdot)) \) are local linear type instead of the local least squares as in Li et al. (2002). Due to the simple structure and simpler bias expression, we focus on \( (\check{\alpha}(\cdot), \check{\beta}(\cdot)) \) to construct the test.

Let \( \delta_0(z, \gamma_0) \) denote a parametric varying coefficient of interest, where \( \delta_0(z, \gamma_0) \) has a known parametric functional form up to a \( k \)-dimensional parameter vector \( \gamma_0 \). Focusing only on the structure of the frontier, we state the null hypothesis as

\[
H_0 : P(\delta(z) = \delta_0(z, \gamma_0)) = 1 \quad \text{for some } \gamma_0 \in \text{a compact subset in } \mathbb{R}^k.
\]

Let \( \hat{\gamma} \) denote a \( \sqrt{n} \)-consistent estimator for \( \gamma^* \) where \( \gamma^* = \gamma_0 \) under \( H_0 \) and \( \delta_0(z) \equiv \delta_0(z, \hat{\gamma}) \). We follow Li et al. (2002) to construct the test statistic based on

\[
\int [\hat{\delta}(z) - \delta_0(z)]' [\hat{\delta}(z) - \delta_0(z)] dz.
\]

Comment 3: Although we use local linear estimator in constructing \( \hat{\delta}(z) \), we can show (a brief sketch of the proof is given in the Appendix) that

\[
I_n(1 + o_p(1)).
\]

Given that \( \hat{\gamma} \) and \( \hat{\theta} \) are both \( \sqrt{n} \)-consistent estimators under the \( H_0 \), \( I_n \) is the same as the test statistic in equation (26) of Li and Racine (2010). Thus, we follow their arguments to construct the modified test statistics as
\[
\hat{I}_n = \frac{1}{n^2 h_1^2} \sum_{i=1}^n \sum_{t=1}^T \sum_{m=1}^n \sum_{\tau=1}^T (1, X'_{it})(1, X'_{m\tau})' \hat{\epsilon}_{it}\hat{\epsilon}_{m\tau} K \left( \frac{\hat{z}_{m\tau} - \hat{z}_{it}}{h_1} \right),
\]
where \(\hat{\epsilon}_{it} = \hat{y}_{it} - (1, X'_{it})\hat{\delta}_0(z_{it})\). The corresponding scaled test statistic is given by \(\hat{T}_n = nT h_1^{3/2} \hat{I}_n / \hat{\sigma}_0\), where
\[
\hat{\sigma}_0^2 = 2 \frac{1}{n^2 h_1^2} \sum_{i=1}^n \sum_{t=1}^T \sum_{m=1}^n \sum_{\tau=1}^T ((1, X'_{it})(1, X'_{m\tau})')^2 (\hat{\epsilon}_{it})^2 (\hat{\epsilon}_{m\tau})^2 K \left( \frac{\hat{z}_{m\tau} - \hat{z}_{it}}{h_1} \right)^2.
\]

Since using asymptotic null distribution of the kernel based consistent model specification tests to conduct hypothesis test can lead to poor size performance, we use a residual-based wild bootstrap method to approximate the null distribution of \(\hat{T}_n\). (i) Specifically, we generate the wild bootstrap error \(\epsilon^*_it\) via a two point distribution such that \(\epsilon^*_it = [(1 - \sqrt{5})/2] \hat{\epsilon}_{it}\) with probability \((1 + \sqrt{5})/(2\sqrt{5})\), \(\epsilon^*_it = [(1 + \sqrt{5})/2] \hat{\epsilon}_{it}\) with probability \((-1 + \sqrt{5})/(2\sqrt{5})\). (ii) From \(\{\epsilon^*_it\}_{i=1}^n\}_{t=1}^T\), we generate the bootstrap sample variable \(y^*it = (1, X'_{it})\delta_0(z_{it}) + \epsilon^*_it\) for \(i = 1, \ldots, n\), and \(t = 1, \ldots, T\). (iii) We use the bootstrap sample \(\{y^*it, X_{it}, z_{it}\}_{i=1}^n\}_{t=1}^T\) to estimate the parametric model under \(H_0\) and obtain \(\hat{\gamma}^*\), i.e., via an Ordinary Least Square estimator if the parametric model is linear. We compute the bootstrap residual \(\hat{\epsilon}_{it}^* = y^*it - (1, X'_{it})\delta_0(z_{it}, \hat{\gamma}^*)\). (iv) The bootstrap test statistic \(\hat{T}_n^*\) is obtained as in \(\hat{T}_n\), with \(\hat{\epsilon}_{it}\hat{\epsilon}_{m\tau}\) replaced by \(\hat{\epsilon}_{it}^*\hat{\epsilon}_{m\tau}^*\). (v) We repeat steps (i)-(iv) for a large number of times, say, \(B = 399\), and we calculate the p-value as \(\frac{1}{B} \sum_{b=1}^B I(\hat{T}_n^*(b) > \hat{T}_n)\), where \(I(\cdot)\) denotes the indicator function and \(\hat{T}_n^*(b)\) refers to the \(b\)-th bootstrap test statistics. In the bootstrap procedure above, we generate the bootstrap sample \(y^*it\) according to the null model, thus the empirical distribution function of \(\hat{T}_n^*\) can be used to approximate the null distribution of \(\hat{T}_n\), regardless of whether the null is valid or not.

In the simulation, we test whether \(z_{it}\) appears in the smooth coefficient frontier. Under the null, \(\delta(z) = \delta_0 = (\alpha_0, \beta_0')'\) such that
\[
y_{it} = \alpha_0 + X'_{it}\beta_0 + v_{it} - u_{it}e^{\alpha z_{it}}.
\]
Specifically, we have \((\alpha_0, \beta_{10}, \beta_{20}) = (-0.6, 4, -1)\) in DGP 1 and \((\alpha_0, \beta_{10}, \beta_{20}) = (1, 2, 3)\) in DGP 2. So we estimate the null model for parameters \((\alpha_0, \beta_{10}, \beta_{20})\) and \((\sigma^2_u, \sigma^2_v, \eta)\) by Maximum Likelihood Estimation (MLE). Under the alternative, we consider the smooth coefficient model as discussed in the last subsection. We generate \(X_{it}\) and \(z_{it}\) as before and fix the parameters \((\sigma^2_u, \sigma^2_v, \eta) = (1, 1, 1)\).

We adopt the same bandwidth and kernel function as in the preceding subsection for estimation of the alternative smooth coefficient frontier model. We report in Table 5 the empirical relative rejection frequency for the bootstrap test statistic proposed above in 500 repetitions, for \(T = 5\) and \(n = (50, 100, 200)\), and for significance levels 0.01, 0.05 and 0.1. We observe that the empirical powers are all 1 for the two alternative models and sample sizes considered, while the empirical sizes for both DGPs under the null
are fairly close to the target level, with the size in GDP 2 being slightly closer to the target relative to the size in DGP 1.

### 1.5 An Empirical Application

In this section, we use the country level aggregate data from the Penn World Table (PWT) to illustrate the applicability of our smooth coefficient frontier function estimators. We extract 2,948 observations in total for 134 countries from 1990 to 2011 and we provide a complete list of countries used in Table 1.6. There are several reasons for choosing this dataset. First, it is by far the most complete country level balanced panel together with variables in labor, capital stock, and human capital per capita. This would allow us to estimate production frontier and evaluate efficiency across a wide range of countries. Second, version 8 of the Penn World Table specifically constructs the output-side real GDP using prices that are constant across different countries and over time, facilitating the comparison of productivity across countries and years.

Our data on output \( Y \) is the real GDP in millions of 2005 US dollars. Input variables are: labor force \( (L) \) and real capital stock \( (K) \), where \( L \) is in millions of persons engaged in employment and \( K \) is in millions of 2005 US dollars. We consider \( H \) as the environment variable, an index of human capital per capita based on years of schooling and returns to education (see Psacharopoulos (1994) and Barro and Lee (2013)). We hypothesize that countries with more human capital tend to produce more output (holding other inputs equal), and this might be due to a change in the marginal product of labor and capital as well as the elasticity of output with respect to labor and capital. We present the summary statistics for the variables \( (\ln(Y), \ln(L), \ln(K) \text{ and } H) \) in Table 1.7 and additional descriptions of the dataset can be obtained in Feenstra et al. (2015).

Allowing the coefficients of \( \ln(L) \) and \( \ln(K) \) to vary flexibly with human capital per capita \( H \), our model (1.1) becomes

\[
\ln(Y_{it}) = \alpha(H_{it}) + \beta_L(H_{it}) \ln(L_{it}) + \beta_K(H_{it}) \ln(K_{it}) + v_{it} - u_i e^{\eta H_{it}}, \tag{1.5}
\]

where \( i = 1, 2, ..., 134 \) and \( t = 1, 2, ..., 22 \). We follow the simulation section to select the bandwidth and kernel function to implement our estimators. \( H \) shifts the production frontier directly via \( \alpha(H_{it}) \), while \( H \) also affects the production frontier indirectly through \( \beta_L(H_{it}) \) and \( \beta_K(H_{it}) \), the output elasticities of labor and capital, respectively. As before, \( v_{it} \) is from \( N(0, \sigma^2_v) \), \( u_i \) from \( |N(0, \sigma^2_u)| \) and \( H \) can affect the efficiency through \( e^{\eta H_{it}} \).

We compare our smooth coefficient model with two benchmark models. The first is a Cobb-Douglas
production function with composed error, viz.,

\[
\ln(Y_{it}) = \alpha_0 + \alpha_1 H_{it} + \alpha_2 H_{it}^2 + \beta_L \ln(L_{it}) + \beta_K \ln(K_{it}) + v_i - u_i \eta H_{it}.
\] (1.6)

We allow the human capital index \( H \) to shift production function neutrally via \( \alpha(H_{it}) = \alpha_0 + \alpha_1 H_{it} \), a quadratic function to allow possible nonlinearity in \( H \), and let the other coefficients \( \beta_L \) and \( \beta_K \) to be constant. It is easy to see that it is nested in our equation (1.5). The second benchmark is a translog model but it is not nested in our equation (1.5) because of the presence of the \( \ln(L_{it})^2 \), \( \ln(K_{it})^2 \) and \( \ln(L_{it}) \ln(K_{it}) \) terms, viz.,

\[
\ln(Y_{it}) = \alpha_0 + \alpha_1 H_{it} + \alpha_2 H_{it}^2 + \beta_{1L} \ln(L_{it}) + \beta_{2L} \ln(L_{it})^2 + \beta_{1K} \ln(K_{it}) + \beta_{2K} \ln(K_{it})^2 + \beta_{HL} H_{it} \ln(L_{it}) + \beta_{HK} H_{it} \ln(K_{it}) + \beta_{LK} \ln(L_{it}) \ln(K_{it}) + v_i - u_i \eta H_{it} \]. (1.7)

We estimate the two benchmark models via the MLE. Given that \( \mu(z_t; \theta_0) = \sigma_u \sqrt{2/\pi} \exp(\eta z_t) \), we present the results below with the standard error in parenthesis where the * on the upper right corner indicates that the parameter is significant at the 1% level.

\[
\begin{align*}
\hat{\ln}(Y_{it}) &= 4.984 \quad -0.094 H_{it} + 0.126 H_{it}^2 + 0.476 \ln(L_{it}) + 0.490 \ln(K_{it}) - \sqrt{293/2/\pi} \exp(0.324 z_{it}) \\
&= (0.213)^* \quad (0.173) \quad (0.034)^* \quad (0.023)^* \quad (0.012)^* \\
\hat{\sigma}_u^2 &= 0.293(0.743)^*, \hat{\sigma}_v^2 = 0.028(0.001)^*, \hat{\eta} = 0.324(0.052)^*, \text{ Log-likelihood } = 3316.77, \text{ sample size } = 2948, \\
\ln(Y_{it}) &= 0.914 \quad +0.692 H_{it} + 0.026 H_{it}^2 + 1.444 \ln(L_{it}) + 0.071 \ln(L_{it})^2 - 0.540 \ln(K_{it}) + 0.057 \ln(K_{it})^2 \\
&= (0.391)^* \quad (0.187)^* \quad (0.149) \quad (0.088)^* \quad (0.012)^* \quad (0.093)^* \quad (0.010)^* \\
&= +0.024 H_{it} \ln(L_{it}) - 0.050 H_{it} \ln(K_{it}) - 0.107 \ln(L_{it}) \ln(K_{it}) - \sqrt{320/2/\pi} \exp(0.211 z_{it}) \\
&= (0.063) \quad (0.071) \quad (0.018)^* \\
\hat{\sigma}_u^2 &= 0.320(0.074)^*, \hat{\sigma}_v^2 = 0.027(0.001)^*, \hat{\eta} = 0.211(0.046)^*, \text{ Log-likelihood } = 3388.54, \text{ sample size } = 2948.
\end{align*}
\]

The estimation results from the Cobb-Douglas model indicate that the coefficients of labor and capital are significantly positive and adds up close to 1. Though the coefficient for \( H \) is insignificant, the coefficient of \( H^2 \) is positive and significant. Therefore, the human capital has a significant direct impact on output.

In the translog model, we observe several insignificant coefficients. Since our semiparametric model would nest equation (1.7) if \( \beta_{2L} = \beta_{2K} = \beta_{LK} = 0 \), we perform a likelihood ratio test for the null of \( \beta_{2L} = \beta_{2K} = \beta_{LK} = 0 \), which gives us a statistic of .0278, smaller than the critical value of 7.82 at 5% significance level. Thus we cannot reject the null hypothesis. After removing the \( \ln(L_{it})^2 \), \( \ln(K_{it})^2 \) and \( \ln(L_{it}) \ln(K_{it}) \) terms from equation (1.7), we estimate the restricted translog model and present the results below.

\[
\begin{align*}
\hat{\ln}(Y_{it}) &= 6.706 \quad -0.392 H_{it} - 0.124 H_{it}^2 + 0.746 \ln(L_{it}) + 0.225 \ln(K_{it}) - 0.128 H_{it} \ln(L_{it}) \\
&= (0.335)^* \quad (0.154) \quad (0.044)^* \quad (0.050)^* \quad (0.039)^* \quad (0.022)^* \\
\end{align*}
\]
Chapter 1. Semiparametric Smooth Coefficient Stochastic Production Frontier

\[ +.126H_{it} \ln(K_{it}) - \sqrt{.442} \sqrt{2\pi} \exp(.217z_{it}) \]

\[ (.018)^* \]

\[ \hat{\sigma}_u^2 = .442(.121)^*, \hat{\sigma}_v^2 = .028(.001)^*, \hat{\eta} = .217(.054)^*, \text{ Log-likelihood } = 3341.79, \text{ sample size } = 2948, \]

The restricted translog model appears to have a much better fit, with all coefficients significant at the 1% level and \( H \) has a p-value of .011. Because of this, in the comparisons below, we focus on the Cobb-Douglas and restricted translog models. Several parametric models show that \( \eta \) is significant, demonstrating the importance of letting the efficiency level to depend on the human capital level. Human capital appears to have significant direct as well as indirect effects on production frontier.

Given the simple structure and bias expression exhibited in \( \hat{\delta}(z) = (\hat{\alpha}(z), \hat{\beta}(z))' \), we focus on it to construct our semiparametric estimates. We present the \( \theta \) parameter estimates, standard errors as well as the 10-th, 90-th percentiles, mean and median of \( \hat{\delta}(z) \) in Table 1.8. We notice that our semiparametric estimates give very different \( \sigma_u^2 \) and \( \sigma_v^2 \) estimates .054 and .037, respectively, compared to .442 and .028 in the restricted translog model, and .293 and .028 in the Cobb-Douglas model. Thus our semiparametric estimates gives a much smaller \( \lambda \) at around 1.46. Given the much larger estimate of \( \eta \ (.531 \text{ instead of } .217 \text{ in the restricted translog model and } .324 \text{ in the Cobb-Douglas model}) \), we believe that human capital plays a much more important role in determining the efficiency level in our semiparametric frontier model. This should translate into a much different picture of efficiency analysis to be performed below.\(^4\)

To give a better picture of the estimated direct effect \( \alpha(\cdot) \) and elasticities of output with respect to labor \( \beta_L(\cdot) \) and capital \( \beta_K(\cdot) \), we plot these estimates against human capital in Figure 1.2 together with their 95% (top panel) and 99% (middle panel) confidence intervals. The Cobb-Douglas and restricted translog estimates are superimposed on the plot for comparison. Although all the impacts are positive, they display very different patterns. Both the parametric models assume a quadratic function for the direct effect \( \alpha(H) \), while the Cobb-Douglas model assumes that the elasticities are independent of human capital, and the restricted translog model specifies that the elasticities are linear in the human capital.

The direct impact \( \alpha(H) \) are decreasing in \( H \) under translog, increasing in \( H \) under Cobb-Douglas. As expected, the elasticities are constant under Cobb-Douglas, while for translog, \( \beta_L(H) \) is decreasing in \( H \) and \( \beta_K(H) \) is increasing in \( H \).

Allowing the direct and indirect effects to be flexible functions of human capital, our estimates of \( \alpha(H) \) and \( \beta_L(H) \) are generally lower than the parametric estimates, while our estimates of \( \beta_K(H) \) are mostly higher, with exceptions only for countries with fairly low level of human capital. The parametric models require that the elasticity is either a constant or a linear function of \( H \), thus most of the effect

\(^4\text{We calculate the standard error of } (\hat{\sigma}_u^2, \hat{\sigma}_v^2, \hat{\eta})' \text{ based on the asymptotic variance expression in Theorem 2. These are } (.0002, .0006, .0147)' \text{ as given in Table 1.8. Although the parameters estimates are significant in all the models, the standard errors of } \hat{\theta} \text{ are much smaller in our semiparametric model.} \)
of human capital on output is only reflected in the intercept term. The semiparametric model allows the elasticity to change flexibly with $H$, thus the direct effect of $H$ on output is much smaller. The confidence bounds for our semiparametric estimates are fairly tight, and the benchmark parametric estimates are outside the confidence bounds for a large portion of $H$, which gives indication that the parametric models are misspecified since the Cobb-Douglas and restricted translog are nested in our semiparametric model.

$\hat{\alpha}(H)$ and $\hat{\beta}_L(H)$ are highly nonlinear in $H$, increasing with $H$ when $H$ is in the middle range from 1.9 to 2.6 and decreasing otherwise, while $\hat{\beta}_K(H)$ is generally increasing in human capital. The fact that $\hat{\beta}_L(H)$ decreases with $H$ for $H$ in the lower or higher range may look puzzling, as we expect that higher level of human capital leads to higher level of labor elasticity of output. We notice that similar observation has been obtained for companies that are state-owned or quasi state-owned in Li et al. (2002), where there is more labor in the companies than the efficient level as government did not allow lay-off to avoid social unrest. Karabarbounis and Neiman (2014) documented a decline of labor share globally over time caused by the decrease in the relative price of investment goods, which induces firms to shift away from labor and toward capital. We sort the countries according to their human capital level in 2011 in Table 1.6, which indicates that countries with $H$ less than 1.9 and higher than 2.6 are mostly developing and developed countries, respectively. Here we utilize the aggregate data and we conjecture that countries with low level of human capital have a surplus of labor and scarcity of capital and more than optimum level of labor is employed in production. For countries with high level of human capital, increased human capital would greatly boost the productivity of capital, which induces firms to employ more capital and less labor thereby reducing the factor share of labor. We also plot in the bottom panel of Figure 1.2 the kernel density estimates of $\hat{\delta}(H)$, which confirm our intuition that $\hat{\alpha}(H)$ and $\hat{\beta}_L(H)$ are more concentrated in the lower range and $\hat{\beta}_K(H)$ more clustered in the higher range.

We plot in the left panel of Figure 1.3 the returns to scale (RTS) defined as the sum of the coefficients on labor and capital for our semiparametric model as well as the two other benchmark parametric models. After increasing sharply with $H$ for $H$ less than 1.6, our semiparametric estimates of RTS remain fairly stable afterwards with values hovering around 0.96. We notice that the average RTS from the semiparametric model (.958) is slightly less than those of the parametric models (0.966). The right panel of Figure 1.3 gives the kernel density plots of RTS from the semiparametric model which confirms our intuition that the RTS is largely concentrated at a value that is slightly less than unity (constant RTS). Our semiparametric estimates that relax the restrictions placed on the functional form of the direct and indirect effects are very different from those of the parametric models. The capital elasticity of output increases with human capital, while the direct effect of human capital and labor elasticity of output from the semiparametric models exhibit nonlinear dependence on human capital and are much
smaller than those in the parametric models. We speculate that the human capital mostly improves the capital elasticity of output in our data set.

Following Jondrow et al. (1982) we derive that the conditional density of \( u_i \) given \( \epsilon_i \) and \( z_i \), which is

\[
f(u_i | \epsilon_i, z_i) = \left[ 1 - \Phi\left( \frac{\sigma_{u_i}}{\sigma_v} \frac{1}{\sqrt{2\pi} \sigma_{u_i}} \right) \right]^{-1} \frac{1}{\sqrt{2\pi} \sigma_{u_i}} \exp\left[ -\frac{1}{2\sigma_{u_i}^2} \left( u_i + \frac{\sigma_{u_i}^2}{\sigma_v^2} T \frac{\epsilon_{it} g(z_{it})}{\text{trunc}_{0}} \right)^2 \right],
\]

which is the density of a normally distributed random variable with mean \( \mu_{u_i} \) and variance \( \sigma_{u_i}^2 \), truncated at zero. So \( E(u_i | \epsilon_i, z_i) = \mu_{u_i} + \frac{\sigma_{u_i} \phi(-\frac{\mu_{u_i}}{\sigma_{u_i}})}{1 - \Phi(-\frac{\mu_{u_i}}{\sigma_{u_i}})} \) and the mode \( M(u_i | \epsilon_i, z_i) = \mu_{u_i} \) if \( T \sum_{t=1}^{T} \epsilon_{it} g(z_{it}) \leq 0 \), and 0 if \( T \sum_{t=1}^{T} \epsilon_{it} g(z_{it}) > 0 \).

We define the technical efficiency as \( TE_i = e^{-M(u_i | \epsilon_i, z_i)} \) and present the kernel densities for the technical efficiency estimates in Figure 1.4. It illustrates that the density of our \( TE_i \) is clustered at a level larger than those of the parametric benchmarks. Recall that our semiparametric estimate of \( \lambda \) is smaller and that of \( \eta \) is larger, respectively, than the parametric ones. It translates into larger predictions on the efficiency levels. For our smooth coefficient model, the mean and median technical efficiency are 0.8289 and 0.8285, respectively. Those for the Cobb-Douglas model are 0.7353 and 0.7284 and for the translog model these are 0.6724 and 0.6477. The densities for \( TE_i \) under the parametric benchmarks are more skewed or centered to the left. For example, in the smooth coefficient model approximately 5% observations have a technical efficiency of 0.7144 or less, while in the Cobb-Douglas model the 5% empirical quantile of the technical efficiency is 0.5727, and in the translog model it is 0.4821. The results suggest that the technical efficiencies are underestimated in the parametric models, perhaps more so in the restricted translog model.

Finally, we performed a model specification test for our semiparametric model against its parametric counterparts using the testing procedures described in the simulation section. The test statistics \( T_n \) for the null of Cobb-Douglas and translog are 312 and 165, respectively. With the boot-strapped \( p \)-value being 0 for both cases, the test strongly rejects the parametric benchmark models. Thus, a quadratic specification for the direct effect of human capital and either a constant or a linear specification for the indirect effect may not be appropriate for the data. Because of this, we prefer the semiparametric specification.

1.6 Conclusion

We propose a smooth coefficient stochastic production frontier model in a panel data framework where distributions of the error and inefficiency components are assumed to be known up to a certain parameter
vector. Both inefficiency and the smooth coefficients are assumed to depend on a vector of exogenous environmental variables that reflect firm characteristics. By specifying the coefficients in the frontier function as unknown smooth functions of the environmental variables the model is made very flexible. Furthermore, the sample size is not as demanding as a fully nonparametric frontier. We propose a multi-step estimation procedure for the smooth coefficients and the parameters associated with noise and inefficiency components, obtain their asymptotic distributions and propose a simple bootstrap test for the frontier functional form specification to conduct inferences. Our simulation results demonstrate good finite sample performances.

We use the aggregated country level panel dataset from the Penn World Table and perform an empirical application of our proposed estimators and the testing procedure. Compared with two benchmark parametric models, our estimates demonstrate that the human capital (the environmental variable in our application) exerts important effects, both direct and indirect (via capital and labor), on the production frontier, and it mostly improves capital elasticity of output. The average efficiencies from the smooth coefficient model are found to be higher than those obtained under two parametric models, which seem to be misspecified in our application.
1.7 Appendix 1: Tables and Figures

Table 1.1: RMSE, BIAS and STDC for parameter estimators across \((\sigma^2, \lambda)\) with \((n, T) = (100, 5)\).

<table>
<thead>
<tr>
<th>Estimators</th>
<th>((\sigma^2, \lambda) = (1.88, 1.66))</th>
<th>((\sigma^2, \lambda) = (1.63, 1.24))</th>
<th>((\sigma^2, \lambda) = (1.35, 0.83))</th>
</tr>
</thead>
<tbody>
<tr>
<td>((\sigma^2_u, \sigma^2_v) = (1.234, .448))</td>
<td>((\sigma^2_u, \sigma^2_v) = (.900, .585))</td>
<td>((\sigma^2_u, \sigma^2_v) = (.517, 0.750))</td>
<td></td>
</tr>
</tbody>
</table>

Panel A:  
\[ y_{it} = -\frac{1}{1+3z_{it}} + x_{1, it}(2z_{it})^3 + x_{2, it}\ln(5z_{it}) + v_{it} - u_ie^{z_{it}} \]

<table>
<thead>
<tr>
<th>Estimators</th>
<th>RMSE</th>
<th>BIAS</th>
<th>STDC</th>
<th>RMSE</th>
<th>BIAS</th>
<th>STDC</th>
<th>RMSE</th>
<th>BIAS</th>
<th>STDC</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\hat{\sigma}^2)</td>
<td>.4616</td>
<td>.2343</td>
<td>.3981</td>
<td>.4194</td>
<td>.1947</td>
<td>.3718</td>
<td>.2482</td>
<td>-.0150</td>
<td>.2480</td>
</tr>
<tr>
<td>(\hat{\lambda})</td>
<td>.3065</td>
<td>-.2377</td>
<td>.1938</td>
<td>.2173</td>
<td>-.1333</td>
<td>.1718</td>
<td>.1809</td>
<td>-.0267</td>
<td>.1791</td>
</tr>
<tr>
<td>(\hat{\eta})</td>
<td>.1668</td>
<td>-.0626</td>
<td>.1547</td>
<td>.2244</td>
<td>-.0737</td>
<td>.2099</td>
<td>.3522</td>
<td>-.1409</td>
<td>.3232</td>
</tr>
<tr>
<td>(\hat{\sigma}^2_u)</td>
<td>.3058</td>
<td>-.0257</td>
<td>.3051</td>
<td>.2864</td>
<td>.0086</td>
<td>.2866</td>
<td>.2896</td>
<td>.0695</td>
<td>.2815</td>
</tr>
<tr>
<td>(\hat{\sigma}^2_v)</td>
<td>.1659</td>
<td>.1642</td>
<td>.0783</td>
<td>.1639</td>
<td>.1427</td>
<td>.0737</td>
<td>.1442</td>
<td>.1183</td>
<td>.0825</td>
</tr>
</tbody>
</table>

Panel B:  
\[ y_{it} = -5\cos(4z_{it}) + x_{1, it}\sin(4z_{it}) + x_{2, it}\ln(\frac{z_{it}}{1-z_{it}}) + v_{it} - u_ie^{z_{it}} \]

<table>
<thead>
<tr>
<th>Estimators</th>
<th>RMSE</th>
<th>BIAS</th>
<th>STDC</th>
<th>RMSE</th>
<th>BIAS</th>
<th>STDC</th>
<th>RMSE</th>
<th>BIAS</th>
<th>STDC</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\hat{\sigma}^2)</td>
<td>.4289</td>
<td>.2364</td>
<td>.3583</td>
<td>.3483</td>
<td>.1762</td>
<td>.3008</td>
<td>.1993</td>
<td>-.0941</td>
<td>.1759</td>
</tr>
<tr>
<td>(\hat{\lambda})</td>
<td>.1888</td>
<td>-.0308</td>
<td>.1865</td>
<td>.1701</td>
<td>.0077</td>
<td>.1702</td>
<td>.1659</td>
<td>.0416</td>
<td>.1608</td>
</tr>
<tr>
<td>(\hat{\eta})</td>
<td>.1214</td>
<td>.0047</td>
<td>.1214</td>
<td>.1673</td>
<td>.0222</td>
<td>.1660</td>
<td>.2423</td>
<td>.0197</td>
<td>.2418</td>
</tr>
<tr>
<td>(\hat{\sigma}^2_u)</td>
<td>.3347</td>
<td>.1414</td>
<td>.3036</td>
<td>.2810</td>
<td>.1090</td>
<td>.2592</td>
<td>.2470</td>
<td>.1018</td>
<td>.2253</td>
</tr>
<tr>
<td>(\hat{\sigma}^2_v)</td>
<td>.0778</td>
<td>.0664</td>
<td>.0406</td>
<td>.0702</td>
<td>.0546</td>
<td>.0441</td>
<td>.0696</td>
<td>.0420</td>
<td>.0555</td>
</tr>
</tbody>
</table>
Table 1.2: RMSE, BIAS and STDC for $\hat{\theta}$ for $(\sigma^2, \lambda) = (1, 1)$ across different sample sizes.

<table>
<thead>
<tr>
<th>$(N, T)$</th>
<th>$\hat{\sigma}_u^2$</th>
<th>RMSE</th>
<th>BIAS</th>
<th>STDC</th>
<th>$\hat{\sigma}_v^2$</th>
<th>RMSE</th>
<th>BIAS</th>
<th>STDC</th>
<th>$\hat{\eta}$</th>
<th>RMSE</th>
<th>BIAS</th>
<th>STDC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(50,5)</td>
<td>.4550</td>
<td>.1036</td>
<td>.4434</td>
<td>.1700</td>
<td>.1354</td>
<td>.1029</td>
<td>.5402</td>
<td>-.2184</td>
<td>.4946</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(100,5)</td>
<td>.2794</td>
<td>.0771</td>
<td>.2689</td>
<td>.1539</td>
<td>.1318</td>
<td>.0796</td>
<td>.3500</td>
<td>-.1725</td>
<td>.3048</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(200,5)</td>
<td>.1951</td>
<td>.0656</td>
<td>.1839</td>
<td>.1311</td>
<td>.1185</td>
<td>.0561</td>
<td>.2780</td>
<td>-.1552</td>
<td>.2309</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(50,10)</td>
<td>.2215</td>
<td>-.0063</td>
<td>.2217</td>
<td>.1784</td>
<td>.1580</td>
<td>.0830</td>
<td>.2887</td>
<td>-.0641</td>
<td>.2818</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(100,10)</td>
<td>.1363</td>
<td>.0071</td>
<td>.1362</td>
<td>.1477</td>
<td>.1375</td>
<td>.0539</td>
<td>.1891</td>
<td>-.0722</td>
<td>.1749</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(200,10)</td>
<td>.0932</td>
<td>.0033</td>
<td>.0932</td>
<td>.1286</td>
<td>.1225</td>
<td>.0393</td>
<td>.1304</td>
<td>-.0573</td>
<td>.1173</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(50,5)</td>
<td>.2703</td>
<td>.1278</td>
<td>.2384</td>
<td>.0831</td>
<td>.0629</td>
<td>.0543</td>
<td>.2890</td>
<td>.0088</td>
<td>.2892</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(100,5)</td>
<td>.1916</td>
<td>.0989</td>
<td>.1643</td>
<td>.0632</td>
<td>.0504</td>
<td>.0381</td>
<td>.1972</td>
<td>.0168</td>
<td>.1967</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(200,5)</td>
<td>.1258</td>
<td>.0633</td>
<td>.1089</td>
<td>.0493</td>
<td>.0412</td>
<td>.0271</td>
<td>.1445</td>
<td>.0326</td>
<td>.1409</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(50,10)</td>
<td>.2470</td>
<td>.1586</td>
<td>.1896</td>
<td>.0836</td>
<td>.0739</td>
<td>.0390</td>
<td>.1871</td>
<td>.0082</td>
<td>.1872</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(100,10)</td>
<td>.1810</td>
<td>.1220</td>
<td>.1338</td>
<td>.0627</td>
<td>.0567</td>
<td>.0267</td>
<td>.1336</td>
<td>.0019</td>
<td>.1337</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(200,10)</td>
<td>.1193</td>
<td>.0821</td>
<td>.0865</td>
<td>.0469</td>
<td>.0433</td>
<td>.0182</td>
<td>.0907</td>
<td>.0161</td>
<td>.0894</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 1.3: Effects of Variance Ratio ($\lambda$) on the BIAS, STDC and RMSE of $\hat{\theta}$ with $(n, T) = (100, 5)$.

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$\hat{\sigma}^2_u$</th>
<th>$\hat{\sigma}^2_v$</th>
<th>$\hat{\eta}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE</td>
<td>BIAS</td>
<td>STDC</td>
</tr>
<tr>
<td>Panel A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.100</td>
<td>.7793</td>
<td>.2016</td>
<td>.7535</td>
</tr>
<tr>
<td>.178</td>
<td>1.110</td>
<td>.3513</td>
<td>1.054</td>
</tr>
<tr>
<td>.316</td>
<td>1.353</td>
<td>.4988</td>
<td>1.259</td>
</tr>
<tr>
<td>.562</td>
<td>.8669</td>
<td>.2485</td>
<td>.8313</td>
</tr>
<tr>
<td>1.000</td>
<td>.2958</td>
<td>.0698</td>
<td>.2878</td>
</tr>
<tr>
<td>1.778</td>
<td>.2466</td>
<td>.0362</td>
<td>.2441</td>
</tr>
<tr>
<td>3.162</td>
<td>.2081</td>
<td>.0087</td>
<td>.2081</td>
</tr>
<tr>
<td>5.623</td>
<td>.2082</td>
<td>-.0005</td>
<td>.2085</td>
</tr>
<tr>
<td>10.000</td>
<td>.1990</td>
<td>-.0035</td>
<td>.1991</td>
</tr>
<tr>
<td>Panel B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$y_{it} = -\frac{1}{1+4z_{it}} + x_{1,it}(2z_{it})^3 + x_{2,it} \ln(5z_{it}) + v_{it} - u_i e^{2z_{it}}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.100</td>
<td>.2597</td>
<td>.0805</td>
<td>.2471</td>
</tr>
<tr>
<td>.178</td>
<td>.2500</td>
<td>.0907</td>
<td>.2332</td>
</tr>
<tr>
<td>.316</td>
<td>.1728</td>
<td>.0823</td>
<td>.1521</td>
</tr>
<tr>
<td>.562</td>
<td>.1605</td>
<td>.0720</td>
<td>.1436</td>
</tr>
<tr>
<td>1.000</td>
<td>.1944</td>
<td>.0989</td>
<td>.1675</td>
</tr>
<tr>
<td>1.778</td>
<td>.2090</td>
<td>.1209</td>
<td>.1707</td>
</tr>
<tr>
<td>3.162</td>
<td>.1963</td>
<td>.1159</td>
<td>.1585</td>
</tr>
<tr>
<td>5.623</td>
<td>.2209</td>
<td>.1413</td>
<td>.1700</td>
</tr>
<tr>
<td>10.000</td>
<td>.2029</td>
<td>.1368</td>
<td>.1500</td>
</tr>
</tbody>
</table>
### Table 1.4: Effects of Sample Size on Smooth Coefficients.

<table>
<thead>
<tr>
<th>(N,T)</th>
<th>RAMSE</th>
<th>ABIAS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{\alpha}$</td>
<td>$\hat{\beta}_1$</td>
</tr>
<tr>
<td>Panel A</td>
<td>$y_{it} = -\frac{1}{1 + e^{4z_{it}}} + x_{1,it}(2z_{it})^3 + x_{2,it} \ln(5z_{it}) + v_{it} - u_{it} e^{\eta z_{it}}$</td>
<td>$\hat{\alpha}$</td>
</tr>
<tr>
<td>(50,5)</td>
<td>.9678</td>
<td>.1020</td>
</tr>
<tr>
<td>(100,5)</td>
<td>.6782</td>
<td>.7016</td>
</tr>
<tr>
<td>(200,5)</td>
<td>.4980</td>
<td>.5182</td>
</tr>
<tr>
<td>(50,10)</td>
<td>.6678</td>
<td>.6988</td>
</tr>
<tr>
<td>(100,10)</td>
<td>.4696</td>
<td>.4850</td>
</tr>
<tr>
<td>(200,10)</td>
<td>.3525</td>
<td>.3681</td>
</tr>
<tr>
<td>Panel B</td>
<td>$y_{it} = -5\cos(4z_{it}) + x_{1,it}\sin(4z_{it}) + x_{2,it} \ln\left(\frac{z_{it}}{10}\right) + v_{it} - u_{it} e^{\eta z_{it}}$</td>
<td>$\hat{\alpha}$</td>
</tr>
<tr>
<td>(50,5)</td>
<td>.9389</td>
<td>.9514</td>
</tr>
<tr>
<td>(100,5)</td>
<td>.6896</td>
<td>.6929</td>
</tr>
<tr>
<td>(200,5)</td>
<td>.5132</td>
<td>.5154</td>
</tr>
<tr>
<td>(50,10)</td>
<td>.7109</td>
<td>.7190</td>
</tr>
<tr>
<td>(100,10)</td>
<td>.5122</td>
<td>.5154</td>
</tr>
<tr>
<td>(200,10)</td>
<td>.3936</td>
<td>.3952</td>
</tr>
</tbody>
</table>
Table 1.5: Empirical relative rejection frequency of the $\hat{T}_n$ statistic with Bootstrap.

<table>
<thead>
<tr>
<th>(n,T)</th>
<th>Empirical Size</th>
<th>Empirical Power</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\alpha = 0.01$</td>
<td>$\alpha = 0.05$</td>
</tr>
<tr>
<td>Panel A</td>
<td>$H_0: y_{it} = -0.6 + 4x_{1,it} + x_{2,it} + v_{it} - u_i e^{z_{it}}$</td>
<td>$H_1: y_{it} = \frac{1}{1+3z_{it}} + x_{1,it}(2z_{it})^3 + x_{2,it} \ln(5z_{it}) + v_{it} - u_i e^{z_{it}}$</td>
</tr>
<tr>
<td>(50,5)</td>
<td>0.0040</td>
<td>0.0300</td>
</tr>
<tr>
<td>(100,5)</td>
<td>0.0020</td>
<td>0.0240</td>
</tr>
<tr>
<td>(200,5)</td>
<td>0.0010</td>
<td>0.0200</td>
</tr>
<tr>
<td>Panel B</td>
<td>$H_0: y_{it} = 1 + 2x_{1,it} + 3x_{2,it} + v_{it} - u_i e^{z_{it}}$</td>
<td>$H_1: y_{it} = -5\cos(4z_{it}) + x_{1,it}\sin(4z_{it}) + x_{2,it} \ln(\frac{z_{it}}{1-2z_{it}}) + v_{it} - u_i e^{z_{it}}$</td>
</tr>
<tr>
<td>(50,5)</td>
<td>0.0035</td>
<td>0.0460</td>
</tr>
<tr>
<td>(100,5)</td>
<td>0.0040</td>
<td>0.0540</td>
</tr>
<tr>
<td>(200,5)</td>
<td>0.0050</td>
<td>0.0480</td>
</tr>
</tbody>
</table>
Table 1.6: List of Countries According to Human Capital (as of 2011)

<table>
<thead>
<tr>
<th>$H \leq 1.9$</th>
<th>$1.9 &lt; H \leq 2.6$</th>
<th>$H &gt; 2.6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benin</td>
<td>Bangladesh</td>
<td>Albania</td>
</tr>
<tr>
<td>Burundi</td>
<td>Brazil</td>
<td>Argentina</td>
</tr>
<tr>
<td>Cambodia</td>
<td>Cameroon</td>
<td>Armenia</td>
</tr>
<tr>
<td>Central African</td>
<td>China</td>
<td>Australia</td>
</tr>
<tr>
<td>Côte d’Ivoire</td>
<td>China: Macao</td>
<td>Austria</td>
</tr>
<tr>
<td>Democratic Congo</td>
<td>Colombia</td>
<td>Bahrain</td>
</tr>
<tr>
<td>Gambia</td>
<td>Congo</td>
<td>Barbados</td>
</tr>
<tr>
<td>Guatemala</td>
<td>Dominican Republic</td>
<td>Belgium</td>
</tr>
<tr>
<td>Malawi</td>
<td>Ecuador</td>
<td>Belgium</td>
</tr>
<tr>
<td>Mali</td>
<td>Egypt</td>
<td>Belize</td>
</tr>
<tr>
<td>Mauritania</td>
<td>El Salvador</td>
<td>Bolivia</td>
</tr>
<tr>
<td>Morocco</td>
<td>Gabon</td>
<td>Botswana</td>
</tr>
<tr>
<td>Mozambique</td>
<td>Ghana</td>
<td>Brunei Darussalam</td>
</tr>
<tr>
<td>Nepal</td>
<td>Honduras</td>
<td>Bulgaria</td>
</tr>
<tr>
<td>Niger</td>
<td>India</td>
<td>Canada</td>
</tr>
<tr>
<td>Rwanda</td>
<td>Indonesia</td>
<td>Chile</td>
</tr>
<tr>
<td>Sierra Leone</td>
<td>Iraq</td>
<td>China: Hong Kong</td>
</tr>
<tr>
<td>Sudan</td>
<td>Kenya</td>
<td>Costa Rica</td>
</tr>
<tr>
<td>Yemen</td>
<td>Kuwait</td>
<td>Croatia</td>
</tr>
<tr>
<td></td>
<td>Lao</td>
<td>Cyprus</td>
</tr>
<tr>
<td></td>
<td>Lesotho</td>
<td>Greek</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$H \leq 1.9$</th>
<th>$1.9 &lt; H \leq 2.6$</th>
<th>$H &gt; 2.6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benin</td>
<td>Bangladesh</td>
<td>Albania</td>
</tr>
<tr>
<td>Burundi</td>
<td>Brazil</td>
<td>Argentina</td>
</tr>
<tr>
<td>Cambodia</td>
<td>Cameroon</td>
<td>Armenia</td>
</tr>
<tr>
<td>Central African</td>
<td>China</td>
<td>Australia</td>
</tr>
<tr>
<td>Côte d’Ivoire</td>
<td>China: Macao</td>
<td>Austria</td>
</tr>
<tr>
<td>Democratic Congo</td>
<td>Colombia</td>
<td>Bahrain</td>
</tr>
<tr>
<td>Gambia</td>
<td>Congo</td>
<td>Barbados</td>
</tr>
<tr>
<td>Guatemala</td>
<td>Dominican Republic</td>
<td>Belgium</td>
</tr>
<tr>
<td>Malawi</td>
<td>Ecuador</td>
<td>Belgium</td>
</tr>
<tr>
<td>Mali</td>
<td>Egypt</td>
<td>Belize</td>
</tr>
<tr>
<td>Mauritania</td>
<td>El Salvador</td>
<td>Bolivia</td>
</tr>
<tr>
<td>Morocco</td>
<td>Gabon</td>
<td>Botswana</td>
</tr>
<tr>
<td>Mozambique</td>
<td>Ghana</td>
<td>Brunei Darussalam</td>
</tr>
<tr>
<td>Nepal</td>
<td>Honduras</td>
<td>Bulgaria</td>
</tr>
<tr>
<td>Niger</td>
<td>India</td>
<td>Canada</td>
</tr>
<tr>
<td>Rwanda</td>
<td>Indonesia</td>
<td>Chile</td>
</tr>
<tr>
<td>Sierra Leone</td>
<td>Iraq</td>
<td>China: Hong Kong</td>
</tr>
<tr>
<td>Sudan</td>
<td>Kenya</td>
<td>Costa Rica</td>
</tr>
<tr>
<td>Yemen</td>
<td>Kuwait</td>
<td>Croatia</td>
</tr>
<tr>
<td></td>
<td>Lao</td>
<td>Cyprus</td>
</tr>
<tr>
<td></td>
<td>Lesotho</td>
<td>Greek</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$H \leq 1.9$</th>
<th>$1.9 &lt; H \leq 2.6$</th>
<th>$H &gt; 2.6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benin</td>
<td>Bangladesh</td>
<td>Albania</td>
</tr>
<tr>
<td>Burundi</td>
<td>Brazil</td>
<td>Argentina</td>
</tr>
<tr>
<td>Cambodia</td>
<td>Cameroon</td>
<td>Armenia</td>
</tr>
<tr>
<td>Central African</td>
<td>China</td>
<td>Australia</td>
</tr>
<tr>
<td>Côte d’Ivoire</td>
<td>China: Macao</td>
<td>Austria</td>
</tr>
<tr>
<td>Democratic Congo</td>
<td>Colombia</td>
<td>Bahrain</td>
</tr>
<tr>
<td>Gambia</td>
<td>Congo</td>
<td>Barbados</td>
</tr>
<tr>
<td>Guatemala</td>
<td>Dominican Republic</td>
<td>Belgium</td>
</tr>
<tr>
<td>Malawi</td>
<td>Ecuador</td>
<td>Belgium</td>
</tr>
<tr>
<td>Mali</td>
<td>Egypt</td>
<td>Belize</td>
</tr>
<tr>
<td>Mauritania</td>
<td>El Salvador</td>
<td>Bolivia</td>
</tr>
<tr>
<td>Morocco</td>
<td>Gabon</td>
<td>Botswana</td>
</tr>
<tr>
<td>Mozambique</td>
<td>Ghana</td>
<td>Brunei Darussalam</td>
</tr>
<tr>
<td>Nepal</td>
<td>Honduras</td>
<td>Bulgaria</td>
</tr>
<tr>
<td>Niger</td>
<td>India</td>
<td>Canada</td>
</tr>
<tr>
<td>Rwanda</td>
<td>Indonesia</td>
<td>Chile</td>
</tr>
<tr>
<td>Sierra Leone</td>
<td>Iraq</td>
<td>China: Hong Kong</td>
</tr>
<tr>
<td>Sudan</td>
<td>Kenya</td>
<td>Costa Rica</td>
</tr>
<tr>
<td>Yemen</td>
<td>Kuwait</td>
<td>Croatia</td>
</tr>
<tr>
<td></td>
<td>Lao</td>
<td>Cyprus</td>
</tr>
<tr>
<td></td>
<td>Lesotho</td>
<td>Greek</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$H \leq 1.9$</th>
<th>$1.9 &lt; H \leq 2.6$</th>
<th>$H &gt; 2.6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benin</td>
<td>Bangladesh</td>
<td>Albania</td>
</tr>
<tr>
<td>Burundi</td>
<td>Brazil</td>
<td>Argentina</td>
</tr>
<tr>
<td>Cambodia</td>
<td>Cameroon</td>
<td>Armenia</td>
</tr>
<tr>
<td>Central African</td>
<td>China</td>
<td>Australia</td>
</tr>
<tr>
<td>Côte d’Ivoire</td>
<td>China: Macao</td>
<td>Austria</td>
</tr>
<tr>
<td>Democratic Congo</td>
<td>Colombia</td>
<td>Bahrain</td>
</tr>
<tr>
<td>Gambia</td>
<td>Congo</td>
<td>Barbados</td>
</tr>
<tr>
<td>Guatemala</td>
<td>Dominican Republic</td>
<td>Belgium</td>
</tr>
<tr>
<td>Malawi</td>
<td>Ecuador</td>
<td>Belgium</td>
</tr>
<tr>
<td>Mali</td>
<td>Egypt</td>
<td>Belize</td>
</tr>
<tr>
<td>Mauritania</td>
<td>El Salvador</td>
<td>Bolivia</td>
</tr>
<tr>
<td>Morocco</td>
<td>Gabon</td>
<td>Botswana</td>
</tr>
<tr>
<td>Mozambique</td>
<td>Ghana</td>
<td>Brunei Darussalam</td>
</tr>
<tr>
<td>Nepal</td>
<td>Honduras</td>
<td>Bulgaria</td>
</tr>
<tr>
<td>Niger</td>
<td>India</td>
<td>Canada</td>
</tr>
<tr>
<td>Rwanda</td>
<td>Indonesia</td>
<td>Chile</td>
</tr>
<tr>
<td>Sierra Leone</td>
<td>Iraq</td>
<td>China: Hong Kong</td>
</tr>
<tr>
<td>Sudan</td>
<td>Kenya</td>
<td>Costa Rica</td>
</tr>
<tr>
<td>Yemen</td>
<td>Kuwait</td>
<td>Croatia</td>
</tr>
<tr>
<td></td>
<td>Lao</td>
<td>Cyprus</td>
</tr>
<tr>
<td></td>
<td>Lesotho</td>
<td>Greek</td>
</tr>
<tr>
<td>Symbol</td>
<td>Variable Name</td>
<td>Mean</td>
</tr>
<tr>
<td>--------</td>
<td>--------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>ln(Y)</td>
<td>Log of Output-side real GDP</td>
<td>10.98</td>
</tr>
<tr>
<td>ln(L)</td>
<td>Log of Labor Force</td>
<td>1.324</td>
</tr>
<tr>
<td>ln(K)</td>
<td>Log of Capital Stock</td>
<td>11.87</td>
</tr>
<tr>
<td>H</td>
<td>Log of Human Capital</td>
<td>2.414</td>
</tr>
</tbody>
</table>
Table 1.8: Estimation of the Semiparametric Smooth Coefficient Stochastic Frontier

<table>
<thead>
<tr>
<th>MLE</th>
<th>Estimates</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\sigma}_u^2$</td>
<td>.0538</td>
<td>.0002</td>
</tr>
<tr>
<td>$\hat{\sigma}_v^2$</td>
<td>.0371</td>
<td>.0006</td>
</tr>
<tr>
<td>$\hat{\eta}$</td>
<td>.5313</td>
<td>.0147</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Smooth Coefficient</th>
<th>10th Percentile</th>
<th>Mean</th>
<th>Median</th>
<th>90th Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\alpha}(H)$</td>
<td>1.220</td>
<td>2.081</td>
<td>1.984</td>
<td>2.509</td>
</tr>
<tr>
<td>$\hat{\beta}_L(H)$</td>
<td>.0703</td>
<td>.1773</td>
<td>.1564</td>
<td>.3070</td>
</tr>
<tr>
<td>$\hat{\beta}_K(H)$</td>
<td>.7090</td>
<td>.7809</td>
<td>.8015</td>
<td>.8875</td>
</tr>
</tbody>
</table>
Figure 1.1: Plots of smooth coefficient functions against $z_{it}$, where $\alpha_1(z_{it}) = -\frac{1}{1+3z_{it}}$, $\beta_{11}(z_{it}) = (2z_{it})^2$, $\beta_{21}(z_{it}) = \ln(5z_{it})$, $\alpha_2(z_{it}) = -5\cos(4z_{it})$, $\beta_{12}(z_{it}) = \sin(4z_{it})$ and $\beta_{22}(z_{it}) = \ln\left(\frac{z_{it}}{1-z_{it}}\right)$.
Figure 1.2: Plots of smooth coefficient frontiers and kernel density estimates: left-\( \alpha(H) \), middle-\( \beta_L(H) \), right-\( \beta_K(H) \). Upper panel: A-smooth coefficient estimates, D and E-95% confidence bound, B-Cobb-Douglas estimate, C-translog estimate. Middle panel: similar to upper, except with 99% confidence bound. Lower panel: kernel density estimates.
Figure 1.3: Plots of Returns to Scale and kernel density estimates

Figure 1.4: Kernel density estimates of technical efficiency
1.8 Appendix 2: Proofs

We denote the local linear regression estimate of $E(\xi|z)$ by $\pi_\xi(z) = \hat{a}_0$ for a generic variable $\xi$, obtained as

$$
(\hat{a}_0, \hat{a}_1) = \arg\min_{a_0, a_1} \sum_{i=1}^n \sum_{t=1}^T (\xi_{it} - a_0 - (z_{it} - z)^T a_1)^2 K \left( \frac{z_{it} - z}{h} \right).
$$

(1.8)

So for $\pi_\mu(z_{it}))$, $\mu_{it}$ refers to $\mu(z_{it}; \theta_0)$, for $\pi_\alpha(z_{it})$, $\alpha_{it}$ refers to $\alpha(z_{it})$, and for $\pi_{X'\beta}(z_{it})$, $(X'\beta)_{it}$ refers to $X'_{it}\beta(z_{it})$.

Lemma 1. $\tilde{\beta}(z) - \beta(z) = \frac{1}{\sum_{j=1}^p \sum_{i=1}^n K(\frac{z_{it} - z}{h})} \left( \hat{s}_{it} - \frac{1}{h} \sum_{j=1}^p \sum_{i=1}^n K(\frac{z_{it} - z}{h}) \right)$ for $l = 1, \ldots, p$, where

$$
\hat{s}_{it} = \frac{1}{2} \sum_{j=1}^p \sum_{i=1}^n \left( \frac{X_{jt} - \beta_j(z)}{h} \right)(X_{jt} - \beta_j(z)) (\xi_{it} - a_0 - (z_{it} - z)^T a_1)^2 K \left( \frac{z_{it} - z}{h} \right).
$$

Proof. We start by recalling Lemma 3 in Lai et al. (1979) that for $\hat{y}_i = \beta_1 X_{it} + \cdots + \beta_p X_{ip} + \epsilon_i$, $X_n = \{X_{i1}, \cdots, X_{ip}\}_{i=1}^n$, $Y_n = \{y_i\}_{i=1}^n$, for the first element $b_{a1}$ in the OLS estimator $b_n = (X_n'X_n)^{-1}X_n'Y_n$, we have

$$
b_{a1} = \beta_1 + \sum_{i=1}^n \sum_{j=1}^p \frac{X_{ij} K_{iT}}{X_n'K_{iT}} T_{ji}.
$$

From equations 1.3 and 1.1, for $\beta_j^{(1)}(z) \equiv \frac{\partial \beta_j(z)}{\partial z}$, $\beta(1)^j(z) = (\beta_1^{(1)}(z), \cdots, \beta_p^{(1)}(z))^T$, we have

$$
\hat{y}_i = y_i - \pi_\mu(z_{it}) = \alpha(z_{it}) + X_{jt} \beta_j(z_{it}) - \mu(z_{it}; \theta_0) + \hat{e}_{it} - \pi_\alpha(z_{it}) - \pi_{X'\beta}(z_{it}) + \pi_\mu(z_{it}) - \pi_\mu(z_{it})
$$

$$
= \sum_{j=1}^p \frac{X_{jt} \beta_j(z) + \beta_j^{(1)}(z)(z_{it} - z) + \hat{e}_{it}}{\sum_{j=1}^p \sum_{i=1}^n K(\frac{z_{it} - z}{h})} + \frac{X_{jt} \beta_j(z)}{\sum_{j=1}^p \sum_{i=1}^n K(\frac{z_{it} - z}{h})} + \hat{e}_{it}.
$$

We obtain the claimed result by applying Lemma 3 in Lai et al. (1979).

Lemma 2. Consider the local linear regression estimate defined in equation 1.8 of $E(\xi|z)$ by $\pi_\xi(z) = \hat{a}_0$ for a generic variable $\xi$. Denote the conditional density of $z$ given $\xi$ by $f_{\xi|z}(z)$. $E(\xi|z) \in C^2$. (4) $nT \alpha \to \infty$, $T$ fixed.

and Assumptions A1, A2, and A3(1), (2), then

$$
(a) \pi_\xi(z) - E(\xi|z) = (1 + o_p(1)) \left\{ \frac{1}{\sum_{i=1}^n \sum_{i=1}^T K(\frac{z_{it} - z}{h})} \right\} \left( \sum_{i=1}^n \sum_{i=1}^T K(\frac{z_{it} - z}{h}) \right) (\xi_{it} - E(\xi|z_{it}))
$$

$$
= (1 + o_p(1)) \left\{ \pi_\xi(z)_{1m} + \pi_\xi(z)_{2m} \right\}.
$$
Chapter 1. Semiparametric Smooth Coefficient Stochastic Production Frontier

(b) $\sup_{z \in \mathbb{G}} |\pi_{z}(z)| = O_p(h^2) + O_p \left( \frac{nh^q}{\ln(n)} \right)^{-\frac{1}{2}}$ and $\sup_{z \in \mathbb{G}} |\pi_{z}(z)| = O_p \left( \frac{nh^q}{\ln(n)} \right)^{-\frac{1}{2}}$, which implies $\sup_{z \in \mathbb{G}} |\pi_{z}(z) - E(\xi|z)| = O_p(L(n))$ for $L(n) = \left( \frac{nh^q}{\ln(n)} \right)^{-\frac{1}{2}} + h^2$.

Proof. (a) Let $S_n(z) = \left[ s_{0n}^{(z)} \ s'_{0n}^{(z)} \right]$ where $s_{0j}^{(z)} = \frac{1}{nh^q} \sum_{i=1}^{n} K \left( \frac{z_i - z_j}{h} \right) z_i$ for $j = 0, 1$, where $z^0 = 1, z^1 = z$, and $s_{02}^{(z)} = \frac{1}{nh^q} \sum_{i=1}^{n} \sum_{j=1}^{T} K \left( \frac{z_i - z_j}{h} \right) (z_i - z_j)$. We further let $w(\psi, z) = (1 0)^T S_n^{-1}(z) (1 \psi)' K(\psi)$. Then

$$
\pi_{z}(z) = (1 0)^T \frac{1}{nh^q} \sum_{i=1}^{n} \sum_{j=1}^{T} S_n^{-1}(z) (1 \ (\frac{z_i - z_j}{h})' K (\frac{z_i - z_j}{h}) \xi_{it}) = \frac{1}{nh^q} \sum_{i=1}^{n} \sum_{j=1}^{T} w(\frac{z_i - z_j}{h}, z) \xi_{it}.
$$

Now we write $\xi_{it} = E(\xi|z_{it}) + \xi_{it} - E(\xi|z_{it})$, and from the discrete moment conditions of the local linear estimation, \( \frac{1}{nh^q} \sum_{i=1}^{n} \sum_{j=1}^{T} w(\frac{z_i - z_j}{h}, z)(z_{it} - z)' = 0 \), \( \frac{1}{nh^q} \sum_{i=1}^{n} \sum_{j=1}^{T} w(\frac{z_i - z_j}{h}, z) = 1 \), so we have

$$
\pi_{z}(z) - E(\xi|z) = \frac{1}{nh^q} \sum_{i=1}^{n} \sum_{j=1}^{T} w(\frac{z_i - z_j}{h}, z)[E(\xi|z_{it}) + \xi_{it} - E(\xi|z_{it}) - E(\xi|z) - E(\xi(z)|z_{it} - z)] = (1 + o_p(1)) \frac{1}{nh^q} \sum_{i=1}^{n} \sum_{j=1}^{T} K \left( \frac{z_i - z_j}{h} \right) \xi_{it}.
$$

(b) With $T$ fixed, we apply Lemma 1 in Yao and Zhang (2015) to obtain the claimed results.

Lemma 3. Define $s_{jn}(z) = \frac{1}{nh^q} \sum_{i=1}^{n} \sum_{j=1}^{T} K \left( \frac{z_i - z_j}{h} \right) \hat{X}_{l, it}^{i}, \hat{X}_{l, it}^{j}$ for $l', l = 1, \ldots, p$, where $z^0 = 1, z^1 = z$, $z^2 = z'$. Then with Assumptions A1, A2, A3(1), (2), (4), A4(5)-(8) and A6(1), we have

(a) $\sup_{z \in \mathbb{G}} |s_{ln}(z) - v_{ll'}(z) f_{l'}(z)| = O_p(L(n))$.

(b) $\sup_{z \in \mathbb{G}} |s_{ln}(z)| = O_p \left( \frac{nh^q}{\ln(n)} \right)^{-\frac{1}{2}} + h^2 1_q$, where $1_q$ is a $q \times 1$ vector of ones.

(c) $\sup_{z \in \mathbb{G}} |s_{ln}(z) - \mu_{l'} z_{it} v_{ll'}(z) f_{l'}(z)| = O_p \left( \frac{nh^q}{\ln(n)} \right)^{-\frac{1}{2}} + o_p(1) = o_p(1)$, where $1_q$ is the identity matrix.

Proof. (a) By Lemma 2, whose assumptions are satisfied by Assumptions A1, A2, A3(1), (2), (4), A 4(5)-(8) and A6(1), we have $\sup_{z \in \mathbb{G}} |s_{ln}(z) - E s_{ln}(z)| = O_p \left( \frac{nh^q}{\ln(n)} \right)^{-\frac{1}{2}}$.

$E s_{ln}(z) = E \left( \frac{1}{nh^q} K \left( \frac{z_i - z_j}{h} \right) v_{ll'}(z_{it}) \right) = \int K(\psi)v_{ll'}(z + h\psi)f_{l'}(z + h\psi) d\psi$.

= $\int K(\psi)[v_{ll'}(z) + h(\frac{\partial v_{ll'}(z)}{\partial \psi}) \psi + \frac{h^2}{2} \psi^2 \frac{\partial^2 v_{ll'}(z)}{\partial \psi^2} \psi] f_{l'}(z) + h(\frac{\partial f_{l'}(z)}{\partial \psi}) \psi + \frac{h^2}{2} \psi^2 \frac{\partial^2 f_{l'}(z)}{\partial \psi^2} \psi] d\psi$.

= $v_{ll'}(z) f_{l'}(z) + O(h^2)$ for $z^* = z + \lambda(z_{it} - z)$ with some $\lambda \in (0, 1)$.

We obtain (b) and (c) with similar arguments.

Theorem 1.

Proof. From Lemma 1, we show

$$
\left( \frac{1}{nh^q} \sum_{i=1}^{n} \sum_{j=1}^{T} K \left( \frac{z_i - z_j}{h} \right) (\hat{X}_{l, it} - K'_{lT} H^{-1}_{TT} T) \right)^2 = f_{l'}(z) \left| V_{ll'}(z) \right| / \left| V_{ll'}(z) \right| + O_p(L(n)).
$$

The second inequality follows from \( \tilde{\alpha}_{i+1} = K_{nT}^{-1}(1 + \tilde{X}_{i+1}H_{nT}^{-1}T_{i+1}) \frac{1}{h} \sum_{j=1}^{p} \tilde{X}_{i+1}(z_{i+1} - z') \beta_j'(z)(z_{i+1} - z) \) 

\[ = (1 + o_p(1)) \frac{1}{h} \int K(\psi) \beta_j'(z) \psi d\psi f(z) |V_{pp}(z)| / |V_p(z)|. \]

(3) \( \sqrt{nT}\tilde{\alpha}_{i} N(0, \int K(\psi) \psi d\psi f(z) V(X_i | z) V(z)) = N(0, fK^{-1}(\psi) \psi d\psi f(z) V(X_i | z) V(z)). \)

(4) \( \sqrt{nT}\tilde{\alpha}_{i} N(0, \int fK^{-1}(\psi) \psi d\psi f(z) V(X_i | z) V(z)). \)

(5) With above results, for \( T_{i+1} \) and \( T_{nT} \) defined below, we write

\[ \sqrt{nT}\tilde{\alpha}_{i} = \frac{1}{nT} \sum_{i=1}^{n} \sum_{j=1}^{T} K(\frac{z_{i+1} - z}{h}) \tilde{X}_{i+1}(\tilde{\alpha}_{i} - K_{nT}^{-1}(1 + \tilde{X}_{i+1}H_{nT}^{-1}T_{i+1}) \tilde{\alpha}_{i} \sqrt{nT}) \]

\[ \tilde{\alpha}_{i} = N(0, fK^{-1}(\psi) \psi d\psi f(z) V(X_i | z) V(z)). \]

So \( \tilde{\alpha}_{i} \rightarrow \tilde{\alpha}_{i} = \beta_i(z) = BI_i(z) + VA_i(z) \) and the claimed results follow. So we show (1)-(4) below.

(1) \( \sqrt{nT} K_{nT} = \frac{1}{nT} \sum_{i=1}^{n} \sum_{j=1}^{T} K(\frac{z_{i+1} - z}{h}) \tilde{X}_{i+1}(\tilde{\alpha}_{i} - (z_{i+1} - z))' = \begin{bmatrix} K_{nT}^{11} \quad K_{nT}^{12} \\ K_{nT}^{21} \quad K_{nT}^{22} \end{bmatrix} \). We denote the element in K_{nT} by K_{nT,j} for \( j = 1, \ldots, p \) and \( j \neq l \).

\[ K_{nT,j} = \frac{1}{nT} \sum_{i=1}^{n} \sum_{j=1}^{T} K(\frac{z_{i+1} - z}{h}) \tilde{X}_{i+1}(\tilde{\alpha}_{i} - (z_{i+1} - z)) \tilde{X}_{j+1}(z_{j+1} - z) + O_p(L(n)) \) from Lemma 2

\[ = \frac{1}{nT} \sum_{i=1}^{n} \sum_{j=1}^{T} K(\frac{z_{i+1} - z}{h}) \tilde{X}_{i+1}(\tilde{\alpha}_{i} - (z_{i+1} - z)) + O_p(L(n)), \]

Since the second inequality follows from \( \tilde{X}_{j+1} = \tilde{X}_{j+1} - (\pi_X(z_{j+1}) - E(X_j | z_{j+1})) \), and we apply Lemma 2 with \( W \) being \( X_j \), together with Assumptions A3(4)-(6), A4(5)-(7), A5, and A6(1).

So \( K_{nT} = \frac{1}{nT} \sum_{i=1}^{n} \sum_{j=1}^{T} K(\frac{z_{i+1} - z}{h}) \tilde{X}_{i+1}S_{i+1} = f_s(z)V_p(z) + O_p(L(n))1_{p-1} \), where \( 1_p = a \times 1 \) vector of ones, and \( V_p(z) = (v_{11}(z), v_{12}(z), \ldots, v_{j+1}(z), v_{j+1}(z), \ldots, v_{p}(z))^' \).

\[ \frac{1}{nT} K_{nT} = \frac{1}{nT} \sum_{i=1}^{n} \sum_{j=1}^{T} K(\frac{z_{i+1} - z}{h}) \tilde{X}_{i+1}S_{i+1} = f_s(z)V_p(z) + O_p(L(n))1_{p-1} \]

since \( \tilde{X}_{j+1} = \tilde{X}_{j+1} + O_p(L(n)) \) and we apply Lemma 3(b).

So we have

\[ \begin{bmatrix} K_{nT}^{11} \\ K_{nT}^{21} \end{bmatrix} = \begin{bmatrix} f_s(z)V_p(z) + O_p(L(n))1_{p-1} \\ O_p(\frac{n\theta^2}{h^2}) + h\theta_{pq} \end{bmatrix}. \]
\[
\begin{align*}
\frac{1}{nTh} H_{nT} &= \frac{1}{nTh} \sum_{i=1}^{n} \sum_{t=1}^{T} K \left( \frac{z_i - z_t}{h} \right) T_{it} T'_{it} \\
&= \frac{1}{nTh} \sum_{i=1}^{n} \sum_{t=1}^{T} K \left( \frac{z_i - z_t}{h} \right) S_{it} S'_{it} \\
&= \begin{bmatrix} H_{0it} & H'_{1it} \\
H_{1it} & H_{2it} \end{bmatrix},
\end{align*}
\]

By Lemma 3, we have \( H_{0it} - V_{zp}(z) f(z) = \sigma_p(L(n)) 1_{p-1} 1_{p-1}' \), \( \frac{1}{h} H_{1it} = \sigma_p \left( \frac{m_{it}}{n_{Tl}} \right)^{1/2} + h \) \( 1_{p-1} 1_{p-1}' \), and \( \frac{1}{h^2} H_{2it} = f_{z}(z) V_{zp}(z) \mu K_{2} L_{q} = \sigma_p(1) 1_{p-1} 1_{p-1}' \). By Assumptions A3(2) and A4(9), \( H^{-1}_{0it} = f_{z}^{-1}(z) V_{zp}^{-1}(z) + \sigma_p(L(n)) 1_{p-1} 1_{p-1}' \) uniformly over \( G \), a compact subset of \( \mathbb{R}^2 \). By the matrix inverse formula, we have
\[
\begin{bmatrix} H_{0it} & H'_{1it} \\
H_{1it} & H_{2it} \end{bmatrix}^{-1} = \begin{bmatrix} H^{-1}_{0it} + H^{-1}_{0it} H'_{1it} F H_{1it} H^{-1}_{1it} & -H^{-1}_{0it} H'_{1it} F \\
-F H_{1it} H^{-1}_{1it} & F \end{bmatrix},
\]
where
\[
h^2 F = [h^2 H_{2it} - (h H_{1it}) H^{-1}_{0it} \frac{1}{h} H'_{1it}]^{-1}
\]
\[
= \left( f_{z}(z) V_{zp}(z) \mu K_{2} L_{q} + \sigma_p(1) 1_{p-1} 1_{p-1}' \right) \times \sigma_p \left( \frac{m_{it}}{n_{Tl}} \right)^{1/2} + h \) \( 1_{p-1} 1_{p-1}' \).
$|V_{pp}(z)|/|V_{pp}(z)|$ when $l = j$.

\begin{equation}
\frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} K \left( \frac{z_{ij} - z}{h} \right) \hat{X}_{i,j} \epsilon_{i,t} = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} K \left( \frac{z_{ij} - z}{h} \right) \left[ \pi_{X_j}(z_i) - E(X_i|z_i) \right] \epsilon_{i,t}.
\end{equation}

With Lemmas 2 and 3, we write

\begin{align*}
& \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} K \left( \frac{z_{ij} - z}{h} \right) \hat{X}_{i,j} \epsilon_{i,t} = \\
& \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} K \left( \frac{z_{ij} - z}{h} \right) \hat{X}_{i,j} \epsilon_{i,t} + \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} K \left( \frac{z_{ij} - z}{h} \right) [- \epsilon_m - \sum_{j=1}^{n} \epsilon_{X_j} \beta_j(\varepsilon_m) + \epsilon_{X_j} \beta_j(\varepsilon_m)](1 + o_p(1))
\end{align*}

When $q = 1$, we have used $\epsilon \in C_0$ as the $q \times q$ second order derivative matrix for a generic function $\xi(z)$, following the notation in Lemma 2. Furthermore, $m_{X_j} \beta_j(\varepsilon_m) \equiv E(X_j|z_j) \beta_j(\varepsilon_m) \equiv X_{j,\varepsilon_m} \beta_j(\varepsilon_m) - E(X_j|\varepsilon_m) \beta_j(\varepsilon_m)$.

Let’s examine $T_{2n}$ first, by ignoring the term of smaller order. Since $T$ is fixed,

\begin{align*}
T_{2n} = & \frac{1}{n^2} \sum_{i=1}^{n} \sum_{m=1}^{n} \sum_{m=1}^{n} \sum_{i=1}^{n} \sum_{m=1}^{n} \sum_{m=1}^{n} \phi_{nim} \left( \frac{z_{ij} - z}{h} \right) K \left( \frac{z_{im} - z}{h} \right) K \left( \frac{z_{jm} - z}{h} \right) \left( \hat{X}_{i,j} \epsilon_{i,t} - \sum_{j=1}^{n} \epsilon_{X_j} \beta_j(\varepsilon_m) \right)
\end{align*}

\begin{align*}
= & \frac{1}{n^2} \sum_{i=1}^{n} \sum_{m=1}^{n} \psi_{nim}
\end{align*}

When $i \neq l$, $T_{2n} = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{m=1}^{n} \sum_{m=1}^{n} \phi_{nim} + \psi_{nim}$ = $\frac{1}{n^2} \sum_{i=1}^{n} \sum_{m=1}^{n} \sum_{m=1}^{n} \phi_{nim}$, since $\phi_{nim}$ is symmetric. We perform the U-statistics projection as in Lemma 1 of Yao and Ullah (2013) to have

$T_{2n} = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{m=1}^{n} \int \phi_{nim} dP(z_m, X_m, \epsilon_m) - \frac{1}{2} E\phi_{nim} + O_p(n^{-1}(E\phi_{nim}^2)^{1/2}).$

Here we define $X_m \equiv (X_{m,1}, \ldots, X_{m,T})'$. We observe that $E(\phi_{nim}) = 0$ and $\int \phi_{nim} dP(z_m, X_m, \epsilon_m) = 0$ since $E(X_{i,j}^2|z_i) = 0$, $E(\epsilon_{i,m}|X_{i,m}, \varepsilon_m) = 0$, and $E(\epsilon_{i,j} \beta_j(z_m)) = 0$.

$E \phi_{nim}^2 \leq CE \phi_{nim}^2$

\begin{align*}
\leq & \frac{1}{n^2} \sum_{i=1}^{n} \sum_{m=1}^{n} \sum_{m=1}^{n} \sum_{m=1}^{n} \phi_{nim} \left( \frac{z_{ij} - z}{h} \right) \left( \frac{z_{im} - z}{h} \right) \left( \frac{z_{jm} - z}{h} \right) \left( \hat{X}_{i,j} \epsilon_{i,t} - \sum_{j=1}^{n} \epsilon_{X_j} \beta_j(\varepsilon_m) \right)
\end{align*}

\begin{align*}
= & O_h(\varepsilon_m^2),
\end{align*}

since $E(\hat{X}_{i,j} \epsilon_{i,t} \gamma_{i,t}^2) = E((X_{i,t} - E(X_{i,t}|z_i)) \gamma_{i,t}) \leq CE(X_{i,t}^2|z_i) \leq C$ as in Assumption A4(6), $E(\varepsilon_m^2|z_m) \leq C[E(\varepsilon_m^2|z_m) + \mu^2(z_m; \theta_0)] \leq C$ by A4(2) and A5(4), and by Assumptions A4(6) and A5(3)

$E(\epsilon_{i,j} \beta_j(z_m)|z_m) = E((X_{i,m} - E(X_{i,m}|z_m)) \beta_j(z_m)|z_m) \leq C \beta_j^2(z_m)E(X_{i,m}^2|z_m) < C$. So we conclude

$T_{2n} = o_P((nh^{-1})^2/2).$
When \( i = l \) and \( t = m \), \( T_{2n} = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{T} \frac{1}{m} K \left( \frac{z_{it} - z_{ij}}{h} \right) \tilde{X}_{i,t}^o \left( - \tilde{e}_{it} - \sum_{j=1}^{p} \epsilon_{X,j} \beta_j \left( z_{it} \right) \right) = O(n^{-1}h^{-q}) \), since \( E(\|\tilde{X}_{i,t}^o \left( - \tilde{e}_{it} - \sum_{j=1}^{p} \epsilon_{X,j} \beta_j \left( z_{it} \right) \right)\|^q) < C \). We can show similarly that for \( i = l \) and \( t \neq m \), \( T_{2n} = O_p(n^{-1}h^{-q}) \).

\[
T_{3n} = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{T} \sum_{k=1}^{T} \frac{1}{m} K \left( \frac{z_{it} - z_{jk}}{h} \right) K \left( \frac{z_{mr} - z_{it}}{h} \right) \tilde{X}_{i,t}^o \times (-1)^j (z_{mr} - z_{it}) [\mu^{(2)}(z_{it}) - \mu^{(2)}(z_{it}, \theta_0)] + \sum_{j=1}^{T} m_{X,j} \beta_j \left( z_{it} \right)(1 + o_p(1))
\]

\[
= \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{T} \psi_{\text{lin}}(1 + o_p(1)) = o_p(n^{-1/2}) + O_p(n^{-1}h^{-q}h^2), \text{ given Assumptions A4(6) and A5.}
\]

Now let’s take a look at \( T_{1n} = \frac{1}{n} \sum_{i=1}^{n} \sum_{t=1}^{T} K \left( \frac{z_{it} - z_{ij}}{h} \right) \tilde{X}_{i,t}^o \tilde{e}_{it} \). By Assumption A4(1), \( E(T_{1n}) = 0 \).

\[
V(T_{1n}) = E(T_{1n}^2) = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{t=1}^{T} E[K \left( \frac{z_{it} - z_{ij}}{h} \right) K \left( \frac{z_{it} - z_{ij}}{h} \right) \tilde{X}_{i,t}^o \tilde{X}_{i,t}^o \tilde{e}_{it} \tilde{e}_{it}]
\]

\[
= \frac{1}{n^2} \sum_{i=1}^{n} \sum_{t=1}^{T} E[K \left( \frac{z_{it} - z_{ij}}{h} \right)^2 \tilde{X}_{i,t}^o \tilde{e}_{it} \tilde{e}_{it}]
\]

\[
= \frac{1}{n} \sum_{i=1}^{n} \sum_{t=1}^{T} E[K \left( \frac{z_{it} - z_{ij}}{h} \right) K \left( \frac{z_{it} - z_{ij}}{h} \right) \tilde{X}_{i,t}^o \tilde{X}_{i,t}^o \tilde{e}_{it} \tilde{e}_{it}]
\]

Note that \( E(\tilde{X}_{i,t}^o | X_{it}, z_{it}) = V(\varepsilon | z_{it}) \), and \( E((\tilde{X}_t^o)^2 | z_{it}) = V(X_i | z_{it}) \), so

\[
nT \delta^2 \sum_{t=1}^{T} E[K \left( \frac{z_{it} - z_{ij}}{h} \right)^2 \tilde{X}_{i,t}^o \tilde{e}_{it} \tilde{e}_{it} = \frac{1}{n^2} E[K \left( \frac{z_{it} - z_{ij}}{h} \right)^2 V(X | z_{it})V(\varepsilon | z)]
\]

\[
\rightarrow \int K^2(\psi) \psi f(z) V(X | z) V(\varepsilon | z)
\]

By the Central Limit Theorem, provided that we can check the Liapunov’s condition, which is

\[
\lim_{n \to \infty} \frac{1}{n^2} \sum_{i=1}^{n} \sum_{t=1}^{T} \frac{1}{m} K \left( \frac{z_{it} - z_{ij}}{h} \right) \tilde{X}_{i,t}^o \tilde{e}_{it} \tilde{e}_{it} = 0 \text{, where } s_n = V(T_{1n}) = O \left( \frac{1}{n^2 \delta^2} \right)
\]

\[
\sqrt{nT \delta^2} T_{1n} \overset{d}{\rightarrow} N(0, \int K^2(\psi) \psi f(z) V(X | z) V(\varepsilon | z)).
\]
Next, we apply Lemma 2 to obtain \( \sup_{z \in \mathcal{G}} |\pi_{X_k}(z) - E(X_k|z)| = O_p(L(n)) \), with Assumptions A3(4) and A4(2), \( \sup_{z \in \mathcal{G}} |\pi(z)| = O_p(L(n)) \). With Assumption A5, \( \sup_{z \in \mathcal{G}} |\pi_{X_k}(z) - \alpha(z)| = O_p(L(n)) \), \( \sup_{z \in \mathcal{G}} |\nu(z) - \mu(z; \theta_0)| = O_p(L(n)) \).

With Assumptions A3(4)-(6), A4(5)-(7), and A5, \( \sup_{z \in \mathcal{G}} |\pi_{X_j,\beta_j}(z) - E(X_j|z)\beta_j(z)| = O_p(L(n)) \). Then we have
\[
\frac{1}{nTh^2} \sum_{i=1}^{n} \sum_{t=1}^{T} K \left( \frac{z_{tt} - z_{it}}{h} \right) \left[ \pi_{X_i}(z_{it}) - E(X_i|z_{it}) \right] \tilde{e}_{it} = K^*_{1i}X_{1i}\tilde{e}_{it} = \sqrt{nTh^3}T_{1n} + o_p(1) \Rightarrow N(0, \int K^2(\psi)d\psi f(z)V(X_i|z)V(\epsilon|z)).
\]

(4) From the results in (1) and (3), we write
\[
K'_{1i}H^{-1}_{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} K \left( \frac{z_{tt} - z_{it}}{h} \right) T_{1i}\tilde{e}_{it} = [V_{1p}(z)V_{i1}(z)^{-1} + O_p(L(n))]V_{p-1,h}^{-1}O_p\left( \frac{nh}{n(nh)} \right)^{-\frac{1}{2}} + O_p(1)\}
\]

\[
\times \left[ \{T_{i1}\Gamma + \alpha_p((nh)^{-1/2})\}^{(n,p)'}_{\Gamma = 1} + O_p((nh)^{-1/2})\Gamma^{(n,p)}_{\Gamma = 1} \right] + o_p(1).
\]

With Assumptions A4(1)-(4) and A4(8)-(10),
\[
nTh^3\text{cov}(T_{1n}, V_{ip}(z)V_{1p}(z)^{-1}T_{1n})_{\Gamma = 1} = V_{1p}(z)V_{ip}(z)^{-1}V_{ip}(z) \int K^2(\psi)d\psi f(x)V(\epsilon|z), \text{ and}
\]

\[
nTh^3V(T_{1n} - V_{ip}(z)V_{ip}(z)^{-1}T_{1n})_{\Gamma = 1} = nTh^3V(T_{1n} - V_{ip}(z)V_{ip}(z)^{-1}T_{1n})_{\Gamma = 1} = \int K^2(\psi)d\psi f(x)V(\epsilon|z)\] \(\Gamma_{\epsilon = 1}^{(n,p)}\).

With the Liapunov’s condition checked as in (3), we have the claimed result.

\[\square\]

**Corollary 1.**

**Proof.** We note that for \( T_{1n} = T_{1n} = \frac{1}{nTh^2} \sum_{i=1}^{n} \sum_{t=1}^{T} K \left( \frac{z_{tt} - z_{it}}{h} \right) X_{1i}\tilde{e}_{it} \),
\[
\hat{\beta}(z) - \beta(z) - B_1(z) = V A_1(z)
\]

\[
= \frac{[V_{1p}(\epsilon)]^{\mathcal{G}}}{\sum_{i=1}^{n} V_{1p}(z)V_{ip}(z)^{-1}T_{1n}}(1 + o_p(1)).
\]
Chapter 1. Semiparametric Smooth Coefficient Stochastic Production Frontier

\[ n \theta_{1h} \text{cov}(VA_i(z), V A_{\ell}(z)) = n \theta_{1h} E \{VA_i(z) V A_{\ell}(z)\} \]

\[ = n \theta_{1h} \mathbb{E}[\psi_i(z) | \psi_{\ell p}(z)] \left\{ E(T_{1n} T_{1n'}) - E(T_{1n} V_{\ell p}(z) V_{\ell p}(z)^{-1} \{T_{1n}\}) \right\}_{k=1, k \neq \ell} \]

\[ = -E(V_{\ell p}(z) V_{\ell p}(z)^{-1} \{T_{1n}\})_{k=1, k \neq \ell} T_{1n'} \]

\[ + E(V_{\ell p}(z) V_{\ell p}(z)^{-1} \{T_{1n}\})_{k=1, k \neq \ell} V_{\ell p}(z) V_{\ell p}(z)^{-1} \{T_{1n}\}) \}

\[ \rightarrow \int K^2(\psi) \mathbb{E}[\psi_i(z) | \psi_{\ell p}(z)] \left\{ E(V_{\ell p}(z) V_{\ell p}(z)^{-1} \{T_{1n}\})_{k=1, k \neq \ell} T_{1n'} \right\} \]

\[ - (k_{\ell p}(z) V_{\ell p}(z)^{-1} \{T_{1n}\})_{k=1, k \neq \ell} + V_{\ell p}(z) V_{\ell p}(z)^{-1} \{T_{1n}\})_{k=1, k \neq \ell} \]

\[ \text{from results (3) and (4) in Theorem 1. With the Liapunov's condition checked as in (3) of Theorem 1, we use the Cramer-Rao device to obtain the claimed result.} \]

\[ \square \]

Theorem 2.

Proof. (1) Let \( \bar{l}_n(\theta, \beta(\cdot), E_{y|x}(\cdot), E_{X|x}(\cdot)) = \frac{n}{\theta} \sum_{i=1}^{n} \ln h(e_i(\theta); z_i, \theta), \) for \( e_i(\theta) = \{e_i(\theta)\}_{i=1}^{n}, e_i(\theta) = \bar{y}_i - \bar{X}_i, \beta(z_i) - \mu(z_i; \theta) \). Let \( \bar{l}_n(\theta, \beta(\cdot), \pi_y(\cdot), \pi_X(\cdot)) = \frac{n}{\theta} \sum_{i=1}^{n} \ln h(e_i(\theta); z_i, \theta) \). Then by definition,

\[ \theta = \arg \max_{\theta \in \Theta} \bar{l}_n(\theta, \beta(\cdot), \pi_y(\cdot), \pi_X(\cdot)) \]

Given Assumptions B1, B2 and the definition of \( \theta \), by Theorem 2.1 in Newey and McFadden (1994), it suffices to prove that \( \sup_{\theta \in \Theta} | \bar{l}_n(\theta, \beta(\cdot), \pi_y(\cdot), \pi_X(\cdot)) - E(\bar{l}_n(\theta, \beta(\cdot), E_{y|x}(\cdot), E_{X|x}(\cdot))) | = o_p(1) \), which follows from

(1) \( \sup_{\theta \in \Theta} | \bar{l}_n(\theta, \beta(\cdot), \pi_y(\cdot), \pi_X(\cdot)) - \bar{l}_n(\theta, \beta(\cdot), E_{y|x}(\cdot), E_{X|x}(\cdot)) | = o_p(1) \)

which follows with Assumption B2, result (3) in Theorem 1 that \( \sup_{x \in \Theta} | E(y|x) | = o_p(1) \), \( \sup_{x \in \Theta} | E(x|x) | = o_p(1) \), and with additional Assumption A3(6), we obtain from Theorem 1 that \( \sup_{x \in \Theta} | \beta(x) - \beta_0(x) | = o_p(1) \).

(2) \( \sup_{\theta \in \Theta} | \bar{l}_n(\theta, \beta(\cdot), E_{y|x}(\cdot), E_{X|x}(\cdot)) - E(\bar{l}_n(\theta, \beta(\cdot), E_{y|x}(\cdot), E_{X|x}(\cdot))) | = o_p(1) \), which we show below.

By the Heine-Borel Theorem, every open covering of \( \Theta \) contains a finite sub-covering \( \{S(\theta_0, d(\theta_0))\}_{k=1}^{K} \), each of which is an open sphere centered at \( \theta_0 \) with radius \( d(\theta_0) > 0 \). Here we did not use \( e_i(\theta) \) since at \( \theta = \theta_0, e \) does not depend on \( \theta \). Then for \( \theta \in S(\theta_0, d(\theta_0)) \),

\[ | \bar{l}_n(\theta, \beta(\cdot), E_{y|x}(\cdot), E_{X|x}(\cdot)) - E(\bar{l}_n(\theta, \beta(\cdot), E_{y|x}(\cdot), E_{X|x}(\cdot))) | = \frac{1}{n} \sum_{i=1}^{n} \ln h(e_i; z_i, \theta) - E \ln h(e_i; z_i, \theta) \]

\[ = \frac{1}{n} \sum_{i=1}^{n} \left( \ln h(e_i; z_i, \theta) - E \ln h(e_i; z_i, \theta) \right) \]

\[ + E \ln h(e_i; z_i, \theta) - E \ln h(e_i; z_i, \theta) \]

\[ \leq \frac{1}{n} \sum_{i=1}^{n} \sup_{\theta \in S(\theta_0, d(\theta_0))} \left| \ln h(e_i; z_i, \theta') - \ln h(e_i; z_i, \theta) \right| - \mu(y_i, x_i, z_i, \theta, \beta(\cdot), E_{y|x}(\cdot), E_{X|x}(\cdot), d(\theta_0)) \]

\[ + E \mu(y_i, x_i, z_i, \theta, \beta(\cdot), E_{y|x}(\cdot), E_{X|x}(\cdot), d(\theta_0)) \]

\[ + \frac{1}{n} \sum_{i=1}^{n} \left( \ln h(e_i; z_i, \theta) - E \ln h(e_i; z_i, \theta) \right) \]

\[ + \left| E \ln h(e_i; z_i, \theta) - E \ln h(e_i; z_i, \theta) \right| \]

where \( \mu(y, x, z, \theta, \beta(\cdot), E_{y|x}(\cdot), E_{X|x}(\cdot), d(\theta_0)) = \sup_{\theta \in S(\theta_0, d(\theta_0))} \left| \ln h(e_i; z_i, \theta') - \ln h(e_i; z_i, \theta) \right| \). By Assumption B1(3), \( \mu(y_i, x_i, z_i, \theta, \beta(\cdot), E_{y|x}(\cdot), E_{X|x}(\cdot), d(\theta)) \rightarrow 0 \) for all \( \theta \in \Theta \) as \( d(\theta) \rightarrow 0 \) almost everywhere. Furthermore, \( \mu(y_i, x_i, z_i, \theta, \beta(\cdot), E_{y|x}(\cdot), E_{X|x}(\cdot), d(\theta)) < 2 \sup_{\theta \in \Theta} \left| \ln h(e_i; z_i, \theta) \right| \), by Assumption B1(4) and the Dominated Convergence Theorem, for any \( \epsilon > 0 \), there exists a \( d > 0 \) such that,
Chapter 1. Semiparametric Smooth Coefficient Stochastic Production Frontier

\( E\mu(y, x, z, \theta, \beta(\cdot), E_{\mid x|}(\cdot), E_{\mid z|}(\cdot), d(\theta)) <\epsilon \) whenever \( d(\theta_k) < d \) for all \( k = 1, \ldots, K \), then

\( E\mu(y, x, z, \theta_k, \beta(\cdot), E_{\mid x|}(\cdot), E_{\mid z|}(\cdot), d(\theta_k)) <\epsilon \) for all \( k \). Also \( |E \ln h(\epsilon_i; z_i, \theta_k) - E \ln h(\epsilon_i; z_i, \theta)| \)

\( \leq E\mu(y, x, z_i, \theta_k, \beta(\cdot), E_{\mid x|}(\cdot), E_{\mid z|}(\cdot), d(\theta_k)) <\epsilon \). So the second and last term above are both less than \( \epsilon \). By the strong Law of Large Numbers, \( E[\ln(\epsilon_i; z_i, \theta_k)] < C \), there exists \( N_{\epsilon,k} \) s.t. \( \forall n > N_{\epsilon,k} \),

\( \frac{1}{n} \sum_{i=1}^{n} \mu(y_i, x, z_i, \theta_k, \beta(\cdot), E_{\mid x|}(\cdot), E_{\mid z|}(\cdot), d(\theta_k)) - E\mu(y, x, z_i, \theta_k, \beta(\cdot), E_{\mid x|}(\cdot), E_{\mid z|}(\cdot), d(\theta_k)) <\epsilon \), and

\( \frac{1}{n} \sum_{i=1}^{n} (\ln h(\epsilon_i; z_i, \theta_k) - E \ln h(\epsilon_i; z_i, \theta)) \) <\( \epsilon \). So above is less than \( 4\epsilon \). Since \( K \) is finite, for all \( n > \max_k N_{\epsilon,k} \), we have the claimed result.

(II) From Equation (1.4), and result in (I) and Assumption B3,

\[ \frac{\partial}{\partial \theta} \bar{I}_n(\theta, \hat{\beta}(\cdot), \pi_y(\cdot), \pi_X(\cdot))|_{\theta=\theta_0}, \]

where \( \bar{I}(\cdot) = I(\theta \in S(\theta_0, d(\theta_0))) \) and \( I(\cdot) \) is the indicator function. We know from (I) that \( \bar{I}(\cdot) \xrightarrow{p} 1 \), so we perform a Taylor expansion to have

\[ (1 + o_p(1)) \frac{\partial^2}{\partial \theta^2} \bar{I}_n(\theta, \hat{\beta}(\cdot), \pi_y(\cdot), \pi_X(\cdot))|_{\theta=\theta_0} = -(1 + o_p(1)) \frac{\partial}{\partial \theta} \bar{I}_n(\theta, \hat{\beta}(\cdot), \pi_y(\cdot), \pi_X(\cdot))|_{\theta=\theta_0}, \]

where for some \( \lambda \in (0, 1) \), \( \theta = \lambda \theta + (1 - \lambda) \theta_0 \). For \( \hat{X}^{\alpha} = (\hat{X}^{\alpha}_{p,i}, \cdots, \hat{X}^{\alpha}_{p,i})' \), we show below that

1. \( \bar{I}_n(\theta, \hat{\beta}(\cdot), \pi_y(\cdot), \pi_X(\cdot)) = \frac{1}{n} \sum_{i=1}^{n} \ln h(\epsilon_i; z_i, \theta) + \frac{1}{n} \sum_{i=1}^{n} \frac{\partial}{\partial \epsilon_i} \ln h(\epsilon_i; z_i, \theta) \]

\[ \times [-(\pi_y(z_{0i}) - E(y|z_{0i})) - \hat{X}^{\alpha}_n(\hat{\beta}(z_{0i}) - \beta(z_{0i})) + \sum_{j=1}^{\rho}(\pi_X_j(z_{0i}) - E(X_j|z_{0i}))\beta_j(z_{0i})] + O_p(L(n)^2) \]

\[ = Q_{1n} + Q_{2n} + O_p(L(n)^2). \]

2. \( \sqrt{n} \left( \frac{\partial^2}{\partial \theta^2} \bar{I}_n(\theta, \hat{\beta}(\cdot), \pi_y(\cdot), \pi_X(\cdot))|_{\theta=\theta_0} - BQ_{2n} \right) \xrightarrow{d} N(0, \sigma^2). \]

3. \( \frac{\partial^2}{\partial \theta^2} \bar{I}_n(\theta, \hat{\beta}(\cdot), \pi_y(\cdot), \pi_X(\cdot))|_{\theta=\theta_0} = \bar{H} + o_p(1). \)

The claim of (II) follows from (1)-(3) above.

1. If \( G \) is a normed linear space and \( T(\hat{g}) : G \to R \) is a functional, we denote the Fréchet differential of \( T \) at \( \hat{g} \) with increment \( \hat{h} \in G \) of order \( i = 1, 2 \) by \( \delta_i^G T(\hat{g}, \hat{h}) \). Note that if the Fréchet differentials of order \( i = 1, 2 \) of \( T \) exist at \( \hat{g} \), they coincide with the Gateaux differentials of order \( i = 1, 2 \) at \( \hat{g} \), denoted by \( \delta_i^G T(\hat{g}; \hat{h}) = \frac{d^i}{d\epsilon^i} T(\hat{g} + \epsilon\hat{h})|_{\epsilon=0} \) (see Luenberger (1969) and Lusternik and Sobolev (1964)). Furthermore, by Taylor's Theorem in Graves (1927),

\( T(\hat{g} + \hat{h}) = T(\hat{g}) + \delta_1^G T(\hat{g}, \hat{h}) + \int_0^1 \delta_1^G T(\hat{g} + \epsilon\hat{h}, \hat{h})(1 - \epsilon) d\epsilon \), where \( \delta_1^G T(\hat{g}, \hat{h}) = \delta_1^G T(\hat{g}, \hat{h}) = \frac{d}{d\epsilon} T(\hat{g} + \epsilon\hat{h})|_{\epsilon=0} \), and

\( \delta_2^G T(\hat{g}, \hat{h}) = \frac{d^2}{d\epsilon^2} T(\hat{g} + \epsilon\hat{h}, \hat{h})|_{\epsilon=0} \), where we take the norm in \( G \) to be the sup norm. Here we let

\( \tilde{g} = (E(y|z_{0i}), E(X|z_{0i})', \beta(z_{0i}))' \),

\( \tilde{h} = ((\pi_y(z_{0i}) - E(y|z_{0i})), \pi_X(z_{0i}) - E(X|z_{0i}))', \cdots, (\pi_y(z_{0i}) - E(Y|z_{0i})), (\hat{\beta}(z_{0i}) - \beta(z_{0i})))' \),

By Taylor's Theorem in Graves (1927),
\[ \hat{l}_n(\theta, \hat{\beta}(\cdot), \pi_y(\cdot), \pi_X(\cdot)) \]
\[ = \frac{1}{n} \sum_{i=1}^{n} \ln h(y_{it} - (E(y|z_{it}) + \pi_y(z_{it}) - E(y|z_{it}))) - \sum_{j=1}^{p} (X_{j, it} - (E(X|z_{it}) + \pi_X(z_{it}) - E(X|z_{it}))) \]
\[ \times (\hat{\beta}_j(z_{it}) + \hat{\beta}_j(z_{it}) - \beta_j(z_{it})) - \mu(z_{it}; \theta)^T \tau_{t=1}^{T} z_{it} \theta \]
\[ = \frac{1}{n} \sum_{i=1}^{n} T(q + \hat{h}_{\tau_{T=1}}^{\tau}) \]
\[ = \frac{1}{n} \sum_{i=1}^{n} \ln h(e_{it}(\theta); z_{it}, \theta) + \frac{1}{n} \sum_{i=1}^{n} \frac{\partial}{\partial \mu \theta} \ln h(e_{it}(\theta); z_{it}, \theta) \]
\[ \times [-\pi(y_{z_{it}} - E(y|z_{it}))) - \hat{X}_{it}^{\prime}(\hat{\beta}(z_{it}) - \beta(z_{it})) + \sum_{j=1}^{p} (\pi_X(z_{it}) - E(X|z_{it}))\hat{\beta}_j(z_{it})] \]
\[ + \frac{1}{n} \sum_{i=1}^{n} \int_0^1 \delta^2_{\tau} T((\hat{g} + r_0 \hat{h}, \hat{h}_{\tau_{T=1}}^{\tau}(1 - r_0))dr_0. \]
\[ \delta^2_{\tau} T((\hat{g} + r_0 \hat{h}, \hat{h}_{\tau_{T=1}}^{\tau(1 - r_0)}) \]
\[ = \frac{1}{n} \sum_{i=1}^{n} \ln h(e_{it}(\theta); z_{it}, \theta) + \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{T} \frac{\partial}{\partial \mu \theta} \ln h(e_{it}(\theta); z_{it}, \theta) \]
\[ \times [-\pi(y_{z_{it}} - E(y|z_{it}))) - \hat{X}_{it}^{\prime}(\hat{\beta}(z_{it}) - \beta(z_{it})) + \sum_{j=1}^{p} (\pi_X(z_{it}) - E(X|z_{it}))\hat{\beta}_j(z_{it})] \]
\[ + \frac{1}{n} \sum_{i=1}^{n} \int_0^1 \delta^2_{\tau} T((\hat{g} + r_0 \hat{h}, \hat{h}_{\tau_{T=1}}^{\tau(1 - r_0)}) \]
\[ \delta^2_{\tau} T((\hat{g} + r_0 \hat{h}, \hat{h}_{\tau_{T=1}}^{\tau(1 - r_0)}) \]
\[ \]
\[ W_{in} = -\pi(y_{z_{it}} - E(y|z_{it}))) - \sum_{j=1}^{p} (\hat{X}_{it}^{\prime} - r_0(\pi_X(z_{it}) - E(X|z_{it}))) \hat{\beta}_j(z_{it}) - \beta_j(z_{it}) \]
\[ + \sum_{j=1}^{p} (\pi_X(z_{it}) - E(X|z_{it}))) \hat{\beta}_j(z_{it}) + r_0(\hat{\beta}_j(z_{it}) - \beta_j(z_{it})) \]
\[ \leq \sup_{y \in G} \pi_y(z) - E(y|z) + \sum_{j=1}^{p} |\hat{X}_{it}^{\prime} - r_0(\pi_X(z_{it}) - E(X|z)))||sup(\hat{\beta}_j(z) - \beta_j(z)) | \]
\[ + \sum_{j=1}^{p} sup_{y \in G} \pi_X(z) - E(X|z))||\hat{\beta}_j(z) + r_0(\hat{\beta}_j(z) - \beta_j(z)) | \]
\[ = O_p(L(n)) + \sum_{j=1}^{p} (|\hat{X}_{it}^{\prime} + \beta_j(z)) + O_p(L(n)) ||\beta_j(z)) + O_p(L(n)) | \]
\[ = O_p(L(n)) + \sum_{j=1}^{p} (|\hat{X}_{it}^{\prime} + \beta_j(z)) + O_p(L(n)) ||\beta_j(z)) + O_p(L(n)) | \]
\[ \]
By Assumptions B3(2) and A5, we obtain \( \frac{1}{n} \sum_{i=1}^{n} \int_0^1 \delta^2_{\tau} T((\hat{g} + r_0 \hat{h}, \hat{h}_{\tau_{T=1}}^{\tau(1 - r_0)}) \]
\[ \]
\[ \]
We observe that \( E(y|z_{it}) = \alpha(z_{it}) + E(X|z_{it}) \beta(z_{it}) - \mu(z_{it}; \theta) = \eta_0(z_{it}) - \mu(z_{it}; \theta) \equiv m(z_{it}; \theta, \eta_0) \) for some \( \theta \in \Theta, \)
\[ \]
\[ \]
\[ \]
\[ \]
with versus the constant true parameter vector $\theta_0$. With Assumption B3, we apply the U-statistics projection in Lemma 1 of Yao and Ulah (2013b) (see also $T_{2n}$ in result (3) of Theorem 1) to obtain

$$Q_{2n1} = \sum_{i=1}^{T} \sum_{t=1}^{T} \frac{\partial}{\partial t} \ln h(\epsilon_i; z_i, \theta)[-(\pi X'_{i}\beta(z_i)) - E(X | z_i')\beta(z_i))]$$

$$= (1 + o_p(1)) \sum_{i=1}^{n} \sum_{t=1}^{T} \sum_{m=1}^{T} \sum_{r=1}^{T} \frac{\partial}{\partial t} \ln h(\epsilon_i; z_i, \theta) \frac{1}{K(\xi_{m,t} - \xi_{r,t})} K(\xi_{m,t} - \xi_{r,t})$$

$$\times \frac{1}{h(\epsilon_m; z_{m,t}, \theta)}h(\epsilon_m; z_{m,t}, \theta_0)de_m$$

$$= Q_{2n11} + Q_{2n12} + o_p(n^{-1/2} + h^2),$$

where $z_{m,t} = (z_{m,t,1}, \ldots, z_{m,t-1,1}, z_{m,t}, z_{m,t+1,1}, \ldots, z_{m,t'})'$ and we have denoted the joint density of $z_i$ by $f_{z_i}(z_i)$.

For $\bar{\epsilon}_{m,t}(\theta) = \epsilon_{m,t} + (\mu(\epsilon_{m,t}, \theta))$, we have

$$Q_{2n2} = \sum_{i=1}^{n} \sum_{t=1}^{T} \frac{\partial}{\partial t} \ln h(\epsilon_i; z_i, \theta)[\epsilon_{m,t}]$$

$$= (1 + o_p(1)) \sum_{i=1}^{n} \sum_{t=1}^{T} \sum_{m=1}^{T} \sum_{r=1}^{T} \frac{\partial}{\partial t} \ln h(\epsilon_i; z_i, \theta) \frac{1}{K(\xi_{m,t} - \xi_{r,t})} K(\xi_{m,t} - \xi_{r,t})$$

$$\times \frac{1}{h(\epsilon_m; z_{m,t}, \theta)}h(\epsilon_m; z_{m,t}, \theta_0)de_m$$

$$= Q_{2n21} + Q_{2n22} + o_p(n^{-1/2})$$

From Theorem 1 result (5), we define $F_D_o(z_{it}) = f_{z_i}(z_i)[V_{zp}(z_{it})]/[V_{zp}(z_{it})] + O_p(L(n))$, then

$$V_Do(z_{it}) = V'_{zp}(z_{it})V_{zp}(z_{it}) + O_p((\frac{m(n)}{n\log n})^{1/2} + h^2)$$

$$S_{r,m} = (\tilde{X}_{1,r,m}, \ldots, \tilde{X}_{n-1,r,m}, \tilde{X}_{n,r,m})'$$

then

$$Q_{2n5} = \sum_{i=1}^{n} \sum_{t=1}^{T} \frac{\partial}{\partial t} \ln h(\epsilon_i; z_i, \theta)(\tilde{X}_{i,1,m} - \tilde{X}_{i,1,m})$$

$$+ \frac{1}{n} \sum_{i=1}^{n} \sum_{t=1}^{T} \frac{\partial}{\partial t} \ln h(\epsilon_i; z_i, \theta)(\tilde{X}_{i,1,m}, \ldots, \tilde{X}_{i,n,r,m})$$

$$= o_p(n^{-1/2}).$$
To obtain the result above, we have used the Lemma 1 in Yao and Ullah (2013b) as in Theorem 1 for the second order U-statistics. We also encountered the third order U-statistics as well and handled it with Theorem 1 in Yao and Martins-Filho (2013) together with Assumption A6(1).

\[
Q_{2n6} = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial}{\partial \tau_{ji}} \ln h(\epsilon_i; z_i, \theta) \sum_{p=1}^{p} \left[ \pi_{X_j}(z_{it}) - E(X_j | z_{it}) \right] \beta_j(z_{it})
\]

\[
= \left( 1 + o_p(1) \right) \frac{\partial}{\partial \tau_{ji}} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{t=1}^{T} \frac{\partial}{\partial \tau_{ji}} \ln h(\epsilon_i; z_i, \theta) \frac{1}{n} \frac{\partial}{\partial f_{x}(z_{it})} K \left( \frac{z_{m, t} - z_{it}}{n} \right) \sum_{j=1}^{p} \beta_j(z_{it})
\]

\[
\times \left[ \frac{1}{3} \left( z_{m, t} - z_{it} \right) E^{(2)}(X_j | z_{it}) (z_{m, t} - z_{it}) + \bar{X}_{j, m, t} \right]
\]

\[
= \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{t=1}^{T} \tilde{X}_{j, m, t} \beta_j(z_{it}) \int \frac{\partial}{\partial \tau_{ji}} \ln h(\epsilon_i; z_i, \theta) h(\epsilon_i; z_i, \theta) d\epsilon_i \int \frac{\partial}{\partial \tau_{ji}} E^{(2)}(X_j | z_{it}) \beta_j(z_{it}) \psi f_{x}(z_{it}) dz_{i} \psi + o_p(h^2)
\]

\[
= Q_{2n61} + Q_{2n62} + o_p(n^{-1/2} + h^2),
\]

So from (1), we can write

\[
\frac{\partial}{\partial \theta_{0}^{}} \left[ Q_{2n12} + Q_{2n21} + Q_{2n31} + Q_{2n62} \right] \theta_{0} = 0
\]

\[
= \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{t=1}^{T} \tilde{X}_{j, m, t} \beta_j(z_{it}) \int \frac{\partial}{\partial \tau_{ji}} \ln h(\epsilon_i; z_i, \theta) h(\epsilon_i; z_i, \theta) d\epsilon_i \int \frac{\partial}{\partial \tau_{ji}} E^{(2)}(X_j | z_{it}) \beta_j(z_{it}) \psi f_{x}(z_{it}) dz_{i} \psi + o_p(h^2)
\]

\[
= Q_{2n11} + Q_{2n12} + Q_{2n21} + Q_{2n31} + Q_{2n62} + o_p(n^{-1/2} + h^2).
\]

We use Assumption B3(4) to interchange order of the differentiation and integration as in Lemma 3.6 of Newey and McFadden (1994) to obtain

\[
\frac{\partial}{\partial \theta_{0}^{}} \left[ Q_{2n12} + Q_{2n21} + Q_{2n31} + Q_{2n62} \right] \theta_{0} = 0
\]

\[
= \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{t=1}^{T} \tilde{X}_{j, m, t} \beta_j(z_{it}) \int \frac{\partial}{\partial \tau_{ji}} \ln h(\epsilon_i; z_i, \theta) h(\epsilon_i; z_i, \theta) d\epsilon_i \int \frac{\partial}{\partial \tau_{ji}} E^{(2)}(X_j | z_{it}) \beta_j(z_{it}) \psi f_{x}(z_{it}) dz_{i} \psi + o_p(h^2)
\]

\[
= Q_{2n11} + Q_{2n12} + Q_{2n21} + Q_{2n31} + Q_{2n62} + o_p(n^{-1/2} + h^2).
\]

We note that \( E(Q_{i}Q_{i}') = 0 \). \( Z_{i} \) is a continuous function of \( \{z_{i}, X_{i}, \epsilon_{i}\} \), which is IID by Assumption A1, by the Cramer-Wald device and Levy’s Central Limit Theorem, \( \sqrt{n} \sum_{i=1}^{n} Q_{i} \rightarrow N(0, \sigma_{F}^2) \), where \( \sigma_{F}^2 = E(Q_{i}Q_{i}') \).

So we obtain

\[
\sqrt{n} \left( \frac{\partial}{\partial \theta_{0}^{}} I_{0}(\theta, \beta(\cdot), \pi_{y}(\cdot), \pi_{X}(\cdot)) \right)_{\theta=\theta_{0}} \rightarrow N(0, \sigma_{F}^2).
\]
Consider the \((a,b)\)th term in \(\frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2}{\partial \theta \partial \theta} \ln h(\epsilon_i; z_i, \tilde{\theta})\). By Assumption B3(1), \(\frac{\partial^2}{\partial \theta \partial \theta} \ln h(\epsilon_i; z_i, \theta)\) is continuous on \(S_{0,\theta}\). We know that \(\tilde{\theta}\) is a consistent estimator of \(\theta_0\). By Assumption B4, \(E\left(\frac{\partial^2}{\partial \theta \partial \theta} \ln h(\epsilon_i; z_i, \theta)\right)\) is continuous at \(\theta\) and 

\[
\sup_{\theta \in S_{0,\theta}} \left| \frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2}{\partial \theta \partial \theta} \ln h(\epsilon_i; z_i, \theta) - E\left(\frac{\partial^2}{\partial \theta \partial \theta} \ln h(\epsilon_i; z_i, \theta)\right) \right| = o_p(1).
\]

By Theorem 21.6 in Davidson (1994), we have 

\[
\frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2}{\partial \theta \partial \theta} \ln h(\epsilon_i; z_i, \tilde{\theta}) \quad \Rightarrow \quad E\left(\frac{\partial^2}{\partial \theta \partial \theta} \ln h(\epsilon_i; z_i, \theta)\right).
\]

With Assumption B4 and similar arguments, we obtain 

\[
\frac{\partial^2}{\partial \theta \partial \theta} [Q_{2n12} + Q_{2n21} + Q_{2n31} + Q_{2n62}||\theta = \tilde{\theta}] = O_p(h^2), \quad \frac{\partial^2}{\partial \theta \partial \theta} [Q_{2n11} + Q_{2n61}||\theta = \theta_0] = o_p(1), \quad \text{and}
\]

\[
\frac{\partial^2}{\partial \theta \partial \theta} [Q_{2n41} + Q_{2n42} + Q_{2n43}||\theta = \theta_0] = (-1)TE\left(\frac{\partial}{\partial \mu} \mu(z_1; \theta_0)\right) \frac{\partial^2}{\partial \theta \partial \theta} \ln h(\epsilon_i; z_i, \theta_0) + \frac{\partial^2}{\partial \theta \partial \theta} \ln h(\epsilon_i; z_i, \theta_0) \frac{\partial}{\partial \theta} \mu(z_i; \theta_0),
\]

which gives the claimed result in (3).

\[\square\]

**Theorem 3.**

**Proof.** From the definition of \(\hat{\alpha}(z)\) and \(\alpha(z) = E(y|z) - \sum_{j=1}^{p} E(X_j|z)\beta_j(z) + \mu(z; \theta_0)\), we have

\[
\hat{\alpha}(z) - \alpha(z) = \pi_y(z; h_1) - E(y|z) - \sum_{j=1}^{p} \left[ (\pi_{X_j}(z; h_1) - E(X_j|z))(\hat{\beta}_j(z; h_1) - \beta_j(z)) \right]
\]

\[
+ \left[ (\pi_{X_j}(z; h_1) - E(X_j|z))\beta_j(z) + E(X_j|z)(\hat{\beta}_j(z; h_1) - \beta_j(z)) \right] + \mu(z; \tilde{\theta}) - \mu(z; \theta_0)
\]

\[
= \pi_{X_j}(z; h_1) - E(X_j|z)\beta_j(z) + \pi_{X_j}(z; h_1) - \alpha(z) - (\pi_{X_j}(z; h_1) - \mu(z; \theta_0)) + \pi_{X_j}(z; h_1)
\]

\[
- \sum_{j=1}^{p} \left[ (\pi_{X_j}(z; h_1) - E(X_j|z))\beta_j(z) + E(X_j|z)(\hat{\beta}_j(z; h_1) - \beta_j(z)) \right] + o_p((nh_1^2)^{-1/2}),
\]

since by Theorem 1 result (5), \( (\pi_{X_j}(z; h_1) - E(X_j|z))(\hat{\beta}_j(z; h_1) - \beta_j(z)) = O_p((n^{-1/2} + h_1^2)^2) = o_p((nh_1^2)^{-1/2}) \) with Assumption A6(2). By Theorem 2, \( \tilde{\theta} - \theta = O_p(n^{-1/2} + h^2) \). With Assumptions B3(3), \( \mu(z; \tilde{\theta}) - \mu(z; \theta_0) = (\frac{\partial \mu(z; \theta)}{\partial \theta})(\tilde{\theta} - \theta_0) = O_p(n^{-1/2} + h^2) = o_p((nh_1^2)^{-1/2}) \), where the undersmoothed \( h \) in Assumption A6(2) has been used to obtain the last equality.

Recall the result in Theorem 1 that \( \hat{\beta}_j(z; h_1) - \beta_j(z) = BI_j(z) + VA_j(z) \), where

\[
BI_j(z) = \frac{h_1^2}{2} \int K(\psi)\psi' \beta_j(\psi) \psi \, d\psi,
\]

\[
VA_j(z) = \frac{\partial \beta_j(z)}{\partial \theta} \mu(z; h_1) - \sum_{j=1}^{p} \left[ (\pi_{X_j}(z; h_1) - E(X_j|z))\beta_j(z) + E(X_j|z)(\hat{\beta}_j(z; h_1) - \beta_j(z)) \right] + o_p((nh_1^2)^{-1/2}).
\]

So we have

\[
\frac{1}{n} \sum_{i=1}^{n} K \left( \frac{z_{i,t} - z}{h_1} \right) \tilde{\epsilon}_{it}(X_{j,t} - V_{j,\theta}(z)V_{\theta}(z)^{-1} S_{ij}).
\]
with Assumption A4(4) that 
\[ E(K(z)|z) = \sum_{i=1}^{p} \sum_{j=1}^{p} E(X_{i,j}^2(z)z^2 + \alpha'(z^2) - \mu'(z^2)z_0)] \]

\[ \times (z_1 - z) \]

\[ + \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} K \left( \frac{z_1 - z}{h_1} \right) \left[ \sum_{j=1}^{p} \sum_{j=1}^{p} E(K_{i,j}(z)|z) + \alpha'(z^2) - \mu'(z^2)z_0) \] \]

By Assumptions A1 and A5,
\[ \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} K \left( \frac{z_1 - z}{h_1} \right) \left[ \sum_{j=1}^{p} \sum_{j=1}^{p} E(K_{i,j}(z)|z) + \alpha'(z^2) - \mu'(z^2)z_0) \] \]

So
\[ A_{1n} + A_{3n} \]

\[ = \frac{k_1}{T} \int K(\psi')\left[ \sum_{j=1}^{p} \sum_{j=1}^{p} E(K_{i,j}(z)|z) + \alpha'(z^2) - \mu'(z^2)z_0) \] \]

\[ \times (z_1 - z) \]

\[ + \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} K \left( \frac{z_1 - z}{h_1} \right) \left[ \sum_{j=1}^{p} \sum_{j=1}^{p} E(K_{i,j}(z)|z) + \alpha'(z^2) - \mu'(z^2)z_0) \] \]

Given 
\[ \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} K \left( \frac{z_1 - z}{h_1} \right) \left[ \sum_{j=1}^{p} \sum_{j=1}^{p} E(K_{i,j}(z)|z) + \alpha'(z^2) - \mu'(z^2)z_0) \] \]

\[ = \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} K \left( \frac{z_1 - z}{h_1} \right) \left[ \sum_{j=1}^{p} \sum_{j=1}^{p} E(K_{i,j}(z)|z) + \alpha'(z^2) - \mu'(z^2)z_0) \] \]

Since 
\[ E(\hat{\varepsilon}_it|X_{i,t}, z_1) = 0 \] and 
\[ E(X_{i,t}^2|z_1) = 0 \] \[ = \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} K \left( \frac{z_1 - z}{h_1} \right) \left[ \sum_{j=1}^{p} \sum_{j=1}^{p} E(K_{i,j}(z)|z) + \alpha'(z^2) - \mu'(z^2)z_0) \] \]

\[ \times (z_1 - z) \]

\[ + \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} K \left( \frac{z_1 - z}{h_1} \right) \left[ \sum_{j=1}^{p} \sum_{j=1}^{p} E(K_{i,j}(z)|z) + \alpha'(z^2) - \mu'(z^2)z_0) \] \]

\[ = \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} K \left( \frac{z_1 - z}{h_1} \right) \left[ \sum_{j=1}^{p} \sum_{j=1}^{p} E(K_{i,j}(z)|z) + \alpha'(z^2) - \mu'(z^2)z_0) \] \]

Theorem 1.4.1. Provided that Assumptions A1, A2, and A5 are satisfied, the smoothed Probit estimator \( \hat{\alpha}(z) = \alpha(z) \) is consistent and asymptotically normal.
For this, we define \( s_n^2 = V(A_{2n} + A_{4n}) = O\left(\frac{1}{n^{2/3}}\right) \), and we need to check
\[
\lim_{n \to \infty} \sum_{i=1}^{n} E\left[ \frac{2}{ \sqrt{n} H(t_i) 1_{\{t_i\} \Gamma(z_i - \tilde{z})} } \right] = 0,
\]
\[
\sum_{i=1}^{n} E\left[ \frac{2}{ \sqrt{n} H(t_i) 1_{\{t_i\} \Gamma(z_i - \tilde{z})} } \right] = \sum_{i=1}^{n} E\left[ \sqrt{n^{2/3}} \frac{2}{ \sqrt{n} H(t_i) 1_{\{t_i\} \Gamma(z_i - \tilde{z})} } \right] \quad |W_{t_i}|^{2+\delta}
\]
\[
\leq C \frac{1}{n^{2/3}} \frac{T}{H(t_i) 1_{\{t_i\} \Gamma(z_i - \tilde{z})}} \leq \frac{1}{n^{2/3}} \frac{T}{H(t_i) 1_{\{t_i\} \Gamma(z_i - \tilde{z})}} |W_{t_i}|^{2+\delta}
\]
\[
\to 0, \quad \text{so the Liapunov’s condition is satisfied with Assumption A4.}
\]

We obtain the claim by combining above results for \( A_{1n}, A_{2n}, A_{3n} \) and \( A_{4n} \).

\( \square \)

**Theorem 4.**

**Proof.** We obtain \( \alpha(z_{it}) = \alpha(z) + \alpha^{(1)}(z)(z_{it} - z) + \frac{1}{2}(z_{it} - z)^2\alpha^{(2)}(z^*)(z_{it} - z) \) with Assumption A5, and similarly \( \beta_j(z_{it}) = \beta_j(z) + \beta_j^{(1)}(z)(z_{it} - z) + \frac{1}{2}(z_{it} - z)^2\beta_j^{(2)}(z^*)(z_{it} - z) \) for \( j = 1, \ldots, p \), where \( z^* = z + \lambda(z_{it} - z) \) for \( \lambda \in (0, 1) \). Then define \( \delta^{(1)}(z) = (\alpha^{(1)}(z), \beta_1^{(1)}(z), \ldots, \beta_p^{(1)}(z))^\prime \), we have
\[
(1, X_{it}^\prime)\delta(z_{it}) = \alpha(z) + \sum_{j=1}^{p} X_{j,it}\beta_j(z_{it} - z) + \frac{1}{2}(z_{it} - z)^2\alpha^{(2)}(z^*)(z_{it} - z)
\]
\[
+ \frac{1}{2}(z_{it} - z)^2\beta_j^{(2)}(z^*)(z_{it} - z) = \hat{Q}_{it}^\prime \left[ \begin{array}{c} \delta(z) \\ \delta^{(1)}(z) \end{array} \right] = \hat{Q}_{it}^\prime \left[ \begin{array}{c} \delta(z) \\ \delta^{(1)}(z) \end{array} \right] + \frac{1}{2}(z_{it} - z)^2[\alpha^{(2)}(z^*) + \sum_{j=1}^{p} X_{j,it}\beta_j^{(2)}(z^*)](z_{it} - z).
\]

Also, \( \mu(z_{it}, \hat{\theta}) - \mu(z_{it}; \theta_0) = \left( \frac{\partial\mu(z_{it}; \hat{\theta})}{\partial \theta} \right) (\hat{\theta} - \theta_0) \). So we have
\[
\hat{\mu}_{it} = (1, X_{it}^\prime)\delta(z_{it}) + \epsilon_{it} + \mu(z_{it}; \theta_0) + \mu(z_{it}; \hat{\theta}) - \mu(z_{it}; \theta_0)
\]
\[
= \hat{Q}_{it}^\prime \left[ \begin{array}{c} \delta(z) \\ \delta^{(1)}(z) \end{array} \right] + \frac{1}{2}(z_{it} - z)^2[\alpha^{(2)}(z^*) + \sum_{j=1}^{p} X_{j,it}\beta_j^{(2)}(z^*)](z_{it} - z) + \epsilon_{it} + \left( \frac{\partial\mu(z_{it}; \hat{\theta})}{\partial \theta} \right) (\hat{\theta} - \theta_0).
\]

Now let \( H = \text{diag}\{I_{p+1}, h_1 I_{(p+1)q}\} \), \( \hat{Q}_{it} = H^{-1}\hat{Q}_{it} = \left[ \begin{array}{c} (1, X_{it}^\prime) \\ (1, X_{it}^\prime) \otimes (\frac{z_{it} - z}{H}) \end{array} \right] \), then
\[
H \left\{ \hat{\alpha} - \left[ \begin{array}{c} \delta(z) \\ \delta^{(1)}(z) \end{array} \right] \right\} = \left[ \sum_{i=1}^{n} K \left( \frac{z_{it} - z}{H} \right) H^{-1}\hat{Q}_{it} \right] \sum_{j=1}^{p} X_{j,it}\beta_j^{(2)}(z^*)(z_{it} - z) + \epsilon_{it} + \left( \frac{\partial\mu(z_{it}; \hat{\theta})}{\partial \theta} \right) (\hat{\theta} - \theta_0)
\]
\[
= \left[ \sum_{i=1}^{n} K \left( \frac{z_{it} - z}{H} \right) H^{-1}\hat{Q}_{it} \right] \sum_{j=1}^{p} X_{j,it}\beta_j^{(2)}(z^*)(z_{it} - z) + \epsilon_{it} + \left( \frac{\partial\mu(z_{it}; \hat{\theta})}{\partial \theta} \right) (\hat{\theta} - \theta_0)
\]
\[
= T_n(z)^{-1} [B_n(z) + V_n(z) + R_n(z)].
\]

(1) Claim: \( T_n(z)^{-1} = T_0(z)^{-1} + o_p(1) \), where \( T_0(z)^{-1} = \left[ \begin{array}{cc} \Omega(z)^{-1} & 0 \\
0 & \frac{\Omega(z)^{-1} \otimes I_q}{T_0(z)} \end{array} \right] \).
\[ \hat{Q}_i \hat{Q}_{it} = \begin{bmatrix} 1 & X_{it} & (\frac{z_{it} - z}{h_1})' & X_{it} \otimes (\frac{z_{it} - z}{h_1})' \\ 1 & X_{it}X_{it}' & X_{it}(\frac{z_{it} - z}{h_1})' & X_{it}X_{it}' \otimes (\frac{z_{it} - z}{h_1})' \\ (\frac{z_{it} - z}{h_1}) & (\frac{z_{it} - z}{h_1})X_{it}' & (\frac{z_{it} - z}{h_1})(\frac{z_{it} - z}{h_1})' & X_{it}' \otimes (\frac{z_{it} - z}{h_1})(\frac{z_{it} - z}{h_1})' \\ X_{it}' \otimes (\frac{z_{it} - z}{h_1})' & X_{it}X_{it}' \otimes (\frac{z_{it} - z}{h_1})(\frac{z_{it} - z}{h_1})' & X_{it}' \otimes (\frac{z_{it} - z}{h_1})' & X_{it}'X_{it}' \otimes (\frac{z_{it} - z}{h_1})' & \end{bmatrix}, \]

we write

\[ T_n(z) = \begin{bmatrix} T_{0n,0}(z) & T_{0n,1}(z) & T_{1n,0}(z) & T_{1n,1}(z) \\ T_{0n,1}(z) & T_{0n,2}(z) & T_{1n,1}(z) & T_{1n,2}(z) \\ T_{1n,0}(z) & T_{1n,1}(z) & T_{2n,0}(z) & T_{2n,1}(z) \\ T_{1n,1}(z) & T_{1n,2}(z) & T_{2n,1}(z) & T_{2n,2}(z) \end{bmatrix}, \]

\[ T_{jn,0}(z) = \frac{1}{n' h_1} \sum_{i=1}^{n} \sum_{j=1}^{T} K \left( \frac{z_{it} - z}{h_1} \right), \]

where \( j = 0, 1, 2 \) and \( z^2 \equiv z^2' \),

\[ T_{jn,1}(z) = \frac{1}{n' h_1} \sum_{i=1}^{n} \sum_{j=1}^{T} K \left( \frac{z_{it} - z}{h_1} \right) X_{it}, \]

and

\[ T_{jn,1}(z) = (T_{jn,11}(z), \ldots, T_{jn,1p}(z))^\prime, \]

\[ T_{jn,1}(z) = (T_{1n,11}(z), \ldots, T_{1n,1p}(z))^\prime, \]

\[ T_{jn,2}(z) = (T_{jn,21}(z), \ldots, T_{jn,2p}(z))^\prime. \]

We apply Lemma 3 on each term and obtain uniformly for all \( z \in \mathcal{G} \) that

\[ T_{0n,0}(z) = f_0(z) + \alpha p(1), \]

\[ T_{1n,0}(z) = \alpha p(1)q', \]

\[ T_{2n,0}(z) = \mu_{K,2} E(X_1'X_1|z) f_0(z) + \alpha p(1)q', \]

\[ T_{1n,1}(z) = \alpha_2 E(X_1'X_1|z) f_0(z) + \alpha p(1)q', \]

\[ T_{jn,1}(z) = \mu_{K,2} E(X_1'X_1|z) f_0(z) + \alpha p(1)q', \]

\[ T_{2n,2}(z) = \mu_{K,2} E(X_1'X_1|z) f_0(z) + \alpha p(1)q', \]

\[ T_{jn,2}(z) = \mu_{K,2} E(X_1'X_1|z) f_0(z) + \alpha p(1)q', \]

So we define

\[ T_0(z) = \begin{bmatrix} \Omega(\beta) & \Omega(\beta) \mu_{K,2} \otimes I_q \\ \Omega(\beta) & \Omega(\beta) \mu_{K,2} \otimes I_q \end{bmatrix} f_0(z), \]

and we have \( T_n(z) = T_0(z) + \alpha p(1)q |(p+1)|x(q+1)|x1 |(p+1)q+1|x1 \) which implies the claim in (1).

(2) Claim: \( T_n(z)^{-1} B_n(z) = h_1^2 \mu_{K,2} (Tr(\phi)^2(\beta)^2))', \]

\[ B_n(z) = \frac{1}{n' h_1^2} \sum_{i=1}^{n} \sum_{j=1}^{T} K \left( \frac{z_{it} - z}{h_1} \right) \begin{bmatrix} 1 & X_{it} \otimes (\frac{z_{it} - z}{h_1})' \\ X_{it} & (\frac{z_{it} - z}{h_1})' \end{bmatrix} \]

\[ \begin{bmatrix} X_{it} \otimes (\frac{z_{it} - z}{h_1})' \\ X_{it} \otimes (\frac{z_{it} - z}{h_1})' \end{bmatrix}, \]

Let \( B_{0n,1}(z) = \frac{1}{n' h_1^2} \sum_{i=1}^{n} \sum_{j=1}^{T} K \left( \frac{z_{it} - z}{h_1} \right) X_{it} \otimes (\frac{z_{it} - z}{h_1})' (\frac{z_{it} - z}{h_1})' \]

\[ B_{0n,1}(z) = \begin{bmatrix} B_{0n,0}(z) \\ B_{0n,1}(z) \\ B_{0n,1}(z) \end{bmatrix}. \]

Furthermore, \( \frac{1}{h_1^2} E B_{0n,1}(z) \rightarrow \frac{1}{2} f_0(z) E[X_1'X_1|z] E(K(z)\psi)' \alpha^{(2)}(z)\psi d\psi + \sum_{j=1}^{p} E(X_1X_1|z) E(K(z)\psi)' \beta^{(2)}_{j}(z)\psi d\psi. \] So we have \( B_{0n,1}(z) = \frac{1}{h_1^2} f_0(z) E(K(z)'E(X_1X_1|z)^2\psi') + \sum_{j=1}^{p} E(X_1X_1|z) E(K(z)' \beta^{(2)}_{j}(z)\psi d\psi + \alpha p(1)q. \) We use similar
arguments on the other terms to obtain
\[ B_n(z) = \frac{n^2}{2} f_z(z) f(\psi) \left( \Omega(z) \psi' \phi(z, \psi) \right) + o_p(h_n)1_{(p+1)(q+1)}, \]
and given Assumption A2 that the kernel function \( K(\cdot) \) is symmetric, we have the claimed result in (2).

(3) Claim: \( \sqrt{nTh_n}T_n^{-1}(z) V_n(z) \xrightarrow{d} N(0, \sigma^2(z)) \), where
\[ S_0(z) = \left( \frac{n}{T} \sum_{t=1}^{T} E \left( \frac{z_{it} - z}{h_n} \right)^2 (\lambda' Q_{it}^2)^2 \right) \frac{1}{nT} \sum_{t=1}^{T} E \left( \frac{z_{it} - z}{h_n} \right)^2 (\lambda' Q_{it}^2)^2 \frac{1}{nT} \sum_{t=1}^{T} E \left( \frac{z_{it} - z}{h_n} \right)^2 (\lambda' Q_{it}^2)^2 \]

Let \( \lambda \) be a \((p+1)(q+1) \times 1\) real vector. Let’s consider \( \lambda' V_n(z) = \frac{n}{T} \sum_{t=1}^{T} \sum_{i=1}^{n} K \left( \frac{z_{it} - z}{h_n} \right)^2 \lambda' Q_{it} \epsilon_{it} \). By Assumption
A4(1), \( E(\lambda' V_n(z)) = 0 \).

\[ V(\lambda' V_n(z)) = \frac{1}{nT} \sum_{t=1}^{T} \sum_{i=1}^{n} E \left( \frac{z_{it} - z}{h_n} \right)^4 (\lambda' Q_{it}^2)^4 + \frac{1}{nT^2} \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{n} E \left( \frac{z_{it} - z}{h_n} \right)^2 \left( \frac{z_{it} - z}{h_n} \right)^2 \left( \frac{z_{jt} - z}{h_n} \right)^4 \right) \]

by the structure of \( \epsilon_{it} \), we can show that
\[ nT \left( \frac{n}{T} \right) \sum_{t=1}^{T} \sum_{i=1}^{n} E \left( \frac{z_{it} - z}{h_n} \right)^2 (\lambda' Q_{it}^2)^4 + \frac{1}{nT^2} \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{n} E \left( \frac{z_{it} - z}{h_n} \right)^2 \left( \frac{z_{it} - z}{h_n} \right)^2 \left( \frac{z_{jt} - z}{h_n} \right)^4 \right) \]

where the equality follows since \( K(\cdot) \) is symmetric, \( K^2(\cdot) \) is also symmetric, so \( \int K(\psi')^2 d\psi = 0 \).

With Assumption A1 and A7(4), we have
\[ nT \left( \frac{n}{T} \right) \sum_{t=1}^{T} \sum_{i=1}^{n} E \left( \frac{z_{it} - z}{h_n} \right)^2 (\lambda' Q_{it}^2)^4 + \frac{1}{nT^2} \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{n} E \left( \frac{z_{it} - z}{h_n} \right)^2 \left( \frac{z_{it} - z}{h_n} \right)^2 \left( \frac{z_{jt} - z}{h_n} \right)^4 \right) \]

Since \( S_0^2 = V(\lambda' V_n(z)) = O((nT \left( \frac{n}{T} \right))^{-1}) \), we apply Assumptions A2, A4(2) and A4(6) to have
\[ \sqrt{nT \left( \frac{n}{T} \right)} \sum_{t=1}^{T} \sum_{i=1}^{n} E \left( \frac{z_{it} - z}{h_n} \right)^2 (\lambda' Q_{it}^2)^4 + \frac{1}{nT^2} \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{n} E \left( \frac{z_{it} - z}{h_n} \right)^2 \left( \frac{z_{it} - z}{h_n} \right)^2 \left( \frac{z_{jt} - z}{h_n} \right)^4 \right) \]

\[ \xrightarrow{d} N(0, \sigma^2(z)) \]

Thus, the Linapov condition is satisfied.

So we conclude that \( \sqrt{nT \left( \frac{n}{T} \right)} T_n^{-1}(z) V_n(z) \xrightarrow{d} N(0, \sigma^2(z) \int \lambda'^2 d\psi) \). The claimed result in (3) follows since
\[ T_n^{-1}(z) \int \lambda'^2 d\psi = S_0(z) \frac{\sigma_n(z)}{\sigma_n(z)^2} \]
(4) Claim: \( T_n(z)^{-1} R_n(z) = o_p((nh^n)^{-1/2}) \).

Given Assumption B3(3) that \( \frac{\partial}{\partial \theta} \mu(z; \theta) < C \), and \( \hat{\theta} - \theta_0 = O_p(n^{-1/2} + h^2) \), by Theorem 2 and Assumption A6(2), we have
\[
|R_n(z)| = \left| \frac{1}{nTH_1} \sum_{i=1}^{n} \sum_{t=1}^{T} K \left( \frac{z_{it} - z}{h} \right) Q_{it} \left( \frac{\partial \mu(z; \hat{\theta})}{\partial \theta} \right)'\left( \hat{\theta} - \theta_0 \right) \right|
\leq \frac{1}{nTH_1} \sum_{i=1}^{n} \sum_{t=1}^{T} K \left( \frac{z_{it} - z}{h} \right) (1, |X_{it}|, |X_{it}'|, |X_{it}'|) \otimes \left( \frac{z_{it} - z}{h} \right)' o_p((nh^n)^{-1/2})
= o_p((nh^n)^{-1/2}) 1_{(p+1)(q+1)},
\]
thus the claimed result follows from (1).

Summarizing the results in (1)-(4), we obtain
\[
\sqrt{nT} \left( \hat{\delta}(z) - \delta(z) \right) \rightarrow N(0, \mathbf{Q}_n(z) \mathbf{Q}_n(z)' + o_p(h^2) 1_{(p+1)(q+1)}) \]
\[
\rightarrow N(0_{(p+1)q+1}, \mathbf{S}_n(z) \mathbf{S}_n(z)' + o_p(h^2) 1_{(p+1)(q+1)}) \]
and the claimed result in Theorem 4 follows.

\[\square\]

A brief proof for the claim in Comment 3:

Proof. Let’s define \( c_{1p} = [I_{p+1}, 0_{(p+1)k}(p+1)] \), where \( I_n \) denotes an identity matrix. Then following notations in Theorem 4, we have
\[
\hat{\delta}(z) = c_{1p}' \left[ \frac{1}{TH_1^2} \sum_{i=1}^{n} \sum_{t=1}^{T} K \left( \frac{z_{it} - z}{h} \right) Q_{it} Q_{it}' \right]^{-1} \frac{1}{TH_1^2} \sum_{i=1}^{n} \sum_{t=1}^{T} K \left( \frac{z_{it} - z}{h} \right) Q_{it} \hat{y}_{it}.
\]
\[
\hat{\delta}_0(z) = c_{1p}' \left[ \frac{1}{TH_1^2} \sum_{i=1}^{n} \sum_{t=1}^{T} K \left( \frac{z_{it} - z}{h} \right) Q_{it} Q_{it}' \right]^{-1} \frac{1}{TH_1^2} \sum_{i=1}^{n} \sum_{t=1}^{T} K \left( \frac{z_{it} - z}{h} \right) Q_{it} \hat{y}_{it} - 1_{(1, X_{it})} \hat{\delta}_0(z)).
\]

So we have
\[
\hat{\delta}(z) - \hat{\delta}_0(z) = c_{1p}' \left[ \frac{1}{TH_1^2} \sum_{i=1}^{n} \sum_{t=1}^{T} K \left( \frac{z_{it} - z}{h} \right) Q_{it} Q_{it}' \right]^{-1} \frac{1}{TH_1^2} \sum_{i=1}^{n} \sum_{t=1}^{T} K \left( \frac{z_{it} - z}{h} \right) Q_{it} \hat{y}_{it} - 1_{(1, X_{it})} \hat{\delta}_0(z)).
\]

Given that \( T_n(z) = \frac{1}{TH_1^2} \sum_{i=1}^{n} \sum_{t=1}^{T} K \left( \frac{z_{it} - z}{h} \right) Q_{it} Q_{it}' \), we observe that the first \((p + 1)\) rows in \( T_n(z)^{-1} \) are \([T_0n(z) - T_1n(z)'T_2n(z)^{-1}T_1n(z)]^{-1}, - (T_0n(z) - T_1n(z)'T_2n(z)^{-1}T_1n(z))^{-1}T_1n(z)'T_2n(z)^{-1}\).

We define \( D_n(z) = T_0n(z) - T_1n(z)'T_2n(z)^{-1}T_1n(z), \) and to avoid the random denominator problem, we insert a positive definite matrix \( D_n(z)'D_n(z) \) between \( [\delta(z) - \hat{\delta}_0(z)]' \) and \( [\hat{\delta}(z) - \hat{\delta}_0(z)] \), so the weighted test statistic becomes \( \int [D_n(z)(\delta(z) - \hat{\delta}_0(z))'] [D_n(z)(\delta(z) - \hat{\delta}_0(z))] dz = I_n + I_{1n}, \) where
\[
I_{1n} = \int \frac{1}{TH_1^2} \sum_{i=1}^{n} \sum_{t=1}^{T} K \left( \frac{z_{it} - z}{h} \right) ((1, X_{it})' \otimes (\frac{z_{it} - z}{h})) (\hat{y}_{it} - (1, X_{it}')) \hat{\delta}_0(z))T_1n(z)T_1n(z)'T_2n(z)^{-1}(z),
\]
\[
\times \frac{1}{TH_1^2} \sum_{i=1}^{n} \sum_{t=1}^{T} K \left( \frac{z_{it} - z}{h} \right) ((1, X_{it}') \otimes (\frac{z_{it} - z}{h})) (\hat{y}_{it} - (1, X_{it}')) \hat{\delta}_0(z)) dz.
\]

Result 1 in the proof of Theorem 4 indicates \( T_2n(z)'T_1n(z)T_1n(z)'T_2n(z)^{-1}(z) = o_p(1) 1_{(p+1)qX(p+1)q}, \) so
\[
I_{1n} \leq C o_p(1) \int \frac{1}{TH_1^2} \sum_{i=1}^{n} \sum_{t=1}^{T} K \left( \frac{z_{it} - z}{h} \right) ((1, X_{it})' \otimes (\frac{z_{it} - z}{h})) (\hat{y}_{it} - (1, X_{it}')) \hat{\delta}_0(z)) dz
\]
\[
\times \frac{1}{TH_1^2} \sum_{i=1}^{n} \sum_{t=1}^{T} K \left( \frac{z_{it} - z}{h} \right) ((1, X_{it}') \otimes (\frac{z_{it} - z}{h})) (\hat{y}_{it} - (1, X_{it}')) \hat{\delta}_0(z)) dz.
\]

We note that \( I_{1n} = O_p(I_n) \) since they only differ in the presence of the term \( (\frac{z_{it} - z}{h}) \), thus we have the claim in Comment 3.

\[\square\]


Chapter 1. Semiparametric Smooth Coefficient Stochastic Production Frontier


Yao, F., Ullah, A., 2013b. Appendix to a nonparametric R-square test for the presence of relevant variables. Economics Department, West Virginia University.

Chapter 2

Does Economic Freedom Affect The Production Frontier?
A Semiparametric Approach With Panel Data

2.1 Introduction

In recent decades, the creation of cross-country measures of economic and political institutions has led to a large literature on the effect of institutions on growth. One measure of economic institutions is the Economic Freedom of the World (EFW) index by Gwartney et al. (2015). The EFW index has been used as a measure of institutions in hundreds of studies in economics and related disciplines (Hall and Lawson, 2014). By far the most frequent relationship of interest to economists has been the impact of institutions on growth. The very first major empirical paper using the EFW index was on this question (Easton and Walker, 1997) and subsequent years have seen dozens of papers written on the economic freedom/growth relationship. Gwartney (2009) summarizes what is known about the relationship between economic freedom and growth in his presidential address to the Southern Economic Association.¹

In this paper we add to this literature by using a smooth coefficient stochastic production frontier model to allow economic freedom to affect a country’s production frontier in a non-linear manner. We

believe that the impact of economic freedom on production frontier cannot be simply captured by entering it into production function linearly as much of the literature assumes. The impact depends further on how economic freedom affects output elasticities (marginal products) of human capital, labor, and physical capital. First, increases in economic freedom improve the mobility of labor and capital across countries (Ashby, 2010; Azman-Saini et al., 2010; Nejad and Young, 2016). For example, an increase in economic freedom for countries with low economic freedom would make physical capital, such as foreign direct investment, more accessible (Bengoa and Sanchez-Robles, 2003; Kapuria-Foreman, 2007). Economic freedom thus alters opportunity costs between labor and capital and affects output elasticities.

Second, economic freedom also reduces transactions costs, which improves productivity in terms of output elasticity (Klein and Luu, 2003).

The effects of economic freedom on output elasticities are largely ignored in the empirical literature on economic freedom and growth. While a handful of studies employ a stochastic production frontier approach to this question, they do not focus on the effect of output elasticities. Instead, their focus is largely on how economic freedom affects technical efficiency. For example, Adkins et al. (2002) investigate the effects of economic freedom on the production frontier and find that more economic freedom in a country leads to decreased inefficiency. They do not, however, address the effect of economic freedom on output elasticities. Similarly, while Klein and Luu (2003) find that increased economic freedom reduces technical inefficiency, they do not allow the elasticity of human or physical capital to vary with the level of institutional quality.

To address this hole in the literature we adopt the semiparametric smooth coefficient model pioneered by Li et al. (2002) and recently extended by Yao et al. (2016). This approach allows us to explicitly explore both the direct effect of economic freedom (the shifting of the production frontier directly caused by economic freedom) and the indirect effect (the shifting of the production frontier due to changing output elasticities). We apply the multi-step procedure developed by Yao et al. (2016) using a smooth coefficient stochastic production frontier to panel data from 1980 to 2010 observed at 5-year increments. We do so in order to estimate the effect of economic freedom ($EF$) on both the production frontiers and technical efficiencies of countries.

To preview our results, we find that increases in economic freedom boost output elasticities of human capital and labor when economic freedom is low or high. Contrary to Hall et al. (2010), we find that the marginal product of human capital and labor is decreasing when countries have an intermediate level of economic freedom. We find, however, that increases in economic freedom improve the marginal product

---

2In this respect, our work is similar to that of Hall et al. (2010), who find in a cross-section that the marginal product of labor and capital varies with the institutional quality of a country.

3Throughout the paper we will use output elasticity and marginal product interchangeably.

4Our empirical approach, however, is different from theirs and thus our results are not directly comparable.
of physical capital when a country has a moderate economic freedom but negatively affects marginal product of physical capital when economic freedom is low or high. The result suggests that for many countries, increases in economic freedom are associated with an outward shift in the production frontier. For example, if China in 2010 had economic freedom at the level of the United States, its production frontier would shift upward by over 20 percent. Finally, we compare the distributions of technical efficiency for our semiparametric model with three parametric counterparts (Cobb-Douglas, translog, and restricted translog) and perform a model specification test which suggest that the semiparametric model is more appropriate for modeling production frontiers.

The remainder of this chapter is organized as following. Section 2 introduces our methodology and model specifications. Section 3 describes the nature of variables and data sources. Section 4 discusses empirical results, illustrating the smooth varying coefficients of stochastic production frontiers and the distribution of technical efficiency. Section 5 concludes.

2.2 Methodology and Model Specification

Adkins et al. (2002) investigate the effects of economic freedom on production frontier and technical efficiency. However, using a parametric Cobb-Douglas production function with composite error, they only examine the direct effect of economic freedom on production frontier but ignore the indirect effects of economic freedom on the marginal productivity of inputs. Klein and Luu (2003) also employ a parametric Cobb-Douglas production function approach. Yao et al. (2016) propose a multi-step estimator for a semiparametric smooth coefficient stochastic production frontier which increases the flexibility of production frontier beyond what a linear or a partially linear model can provide.

Consider the following production frontier with panel data, for \(i = 1, \ldots, n\) and \(t = 1, \ldots, T\),

\[
Y_{it} = \alpha(\text{EF}_{it}) + \beta_H(\text{EF}_{it})\text{Ln}(H_{it}) + \beta_L(\text{EF}_{it})\text{Ln}(L_{it}) + \beta_K(\text{EF}_{it})\text{Ln}(K_{it}) + \epsilon_{it} \\
= (1, X'_{it})(\alpha(\text{EF}_{it}), \beta(\text{EF}_{it})')' + \epsilon_{it} \\
= (1, X'_{it})\delta(\text{EF}_{it}) + \epsilon_{it} \quad (2.1)
\]

where \(Y_{it}\) is the logarithm of output, \(X_{it} = (\text{Ln}(H_{it}), \text{Ln}(L_{it}), \text{Ln}(K_{it}))'\) represent the logarithm of traditional inputs, including human capital, labor and capital, and \(\delta(\text{EF}_{it}) = (\alpha(\text{EF}_{it}), \beta(\text{EF}_{it})')'\), is a vector of unknown smooth functions of exogenous environmental variable, the economic freedom. The composite error is \(\epsilon_{it} = v_{it} - u_{it}\), where we specify a two-sided error \(v_{it} \sim i.i.d.N(0, \sigma_v^2)\) representing random noise, \(u_{it} = u_i g(\text{EF}_{it}; \eta)\) for a one-sided error \(u_i\), scaled by a nonnegative function \(g(\text{EF}_{it}; \eta)\) known up to a parameter \(\eta\), capturing ineffi-
ciency. Here we consider \( u_i \sim i.i.d.|N(0, \sigma^2_u) |\) and independent of \( EF_{it} \) and \( X_{it} \). \( EF_{it} \) enters inefficiency term through the scaling function \( g(\cdot) \) to affect the distribution of \( \epsilon_{it} \). For \( EF_i = (EF_{i1}, \ldots, EF_{iT})' \) and \( \epsilon_i = (\epsilon_{i1}, \ldots, \epsilon_{iT})' \), we denote the conditional density of \( \epsilon_i \) given \( EF_i \) by \( h(\epsilon_i; EF_i, \theta_0) \), where \( \theta_0 = (\sigma^2_u, \sigma^2_v, \eta)' \) denote the true parameters. Thus with above distribution specifications, \( E(\epsilon_{it}|EF_{it}) = -\mu(EF_{it}; \theta_0) = -\sqrt{2 \pi} \sigma_u g(EF_{it}, \eta) \) and \( V(\epsilon_{it}|EF_{it}) = \sigma^2_v + \frac{\pi - 2}{\pi} \sigma^2_u g^2(EF_{it}, \eta) \). We employ a Gaussian kernel function \( K(u) = e^{-u^2/2}/\sqrt{2\pi} \) and a simple rule-of-thumb bandwidth of \( h = 1.06\sigma_{EF}^{-1/5} \) to obtain \( \hat{\theta}(EF_{it}) \), where \( \sigma_{EF} \) represents the standard deviation of economic freedom.

Since \( E(\epsilon_{it}|EF_{it}) \neq 0 \), the standard smooth varying coefficient estimation as in Li et al. (2002) cannot be applied directly. Instead, subtracting conditional mean of Equation (2.1) on both sides, we have

\[
Y_{it} - E(Y_{it}|EF_{it}) = (X_{it} - E(X_{it}|EF_{it}))'\beta(EF_{it}) + \mu(EF_{it}; \theta_0) + \epsilon_{it}, \tag{2.2}
\]

then we estimate \( \hat{\beta}(EF_{it}) \) and \( \hat{\theta} \) with the following steps (See Yao et al. (2016)\(^5\)).

First, let \( \bar{\epsilon}_{it} = \epsilon_{it} + \mu(EF_{it}; \theta_0) \), then \( E(\bar{\epsilon}_{it}|EF_{it}) = 0 \). From Equation (2.2), we construct \( \bar{Y}_{it} = Y_{it} - \bar{\hat{E}}(Y_{it}|EF_{it}), \bar{X}_{it} = X_{it} - \bar{\hat{E}}(X_{it}|EF_{it}) \), where \( \bar{\hat{E}}(Y_{it}|EF_{it}) \) and \( \bar{\hat{E}}(X_{it}|EF_{it}) \) are local linear estimates of conditional mean of \( Y_{it} \) and \( X_{it} \), respectively, evaluated at \( EF_{it} \). Then, Equation (2.2) transforms into,

\[
\bar{y}_{it} = \bar{X}_{it}'\hat{\beta}(EF_{it}) + \bar{\epsilon}_{it}, \tag{2.3}
\]

and we apply standard smooth varying coefficient estimation (see Li et al. (2002)) on Equation (2.3) to obtain consistent estimator \( \hat{\beta}(EF) = (\hat{\beta}_L(EF), \hat{\beta}_H(EF), \hat{\beta}_K(EF))' \), \( \hat{\beta}_1 = (\hat{\beta}_{1L}, \hat{\beta}_{1L}, \hat{\beta}_{1K})' \), \( \hat{\beta}_0 = (\hat{\beta}_0H, \hat{\beta}_0L, \hat{\beta}_0K)' \), \( \hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1)' \), such that

\[
\hat{\beta} = \text{argmin}_b \sum_{i=1}^n \sum_{t=1}^T K \left( \frac{EF_{it} - EF}{h} \right) (\bar{y}_{it} - \bar{W}_{it} b)^2, \tag{2.4}
\]

where \( \bar{W}_{it} = \begin{bmatrix} \bar{X}_{it} \\ \bar{X}_{it} \otimes (EF_{it} - EF) \end{bmatrix} \), and \( \otimes \) denotes the Kronecker product.

Second, recall that \( \mu(EF_{it}; \theta_0) \) is known up to the parameter \( \theta_0 \), we can construct \( \bar{\epsilon}_{it}(\theta) = \bar{y}_{it} - \bar{X}_{it}\bar{\hat{\beta}}(EF_{it}) - \mu(EF_{it}; \theta) \), and \( \hat{\epsilon}(\theta) = \{\bar{\epsilon}_{it}(\theta)\}_{t=1}^T \). We estimate \( \theta \) by \( \hat{\theta} \) via pseudo-likelihood estimation.

\(^5\)Yao et al. (2016) propose a four-step semiparametric estimators, establish their consistency and asymptotic normality, and carry out a comprehensive Monte Carlo study.
Following Pitt and Lee (1981), we write the log-likelihood function as,

\[ \ln \prod_{i=1}^{n} h(\hat{c}_i(\theta); EF \_it, \theta) = \] 

\[ C - \frac{n(T-1)}{2} \ln \sigma_v^2 - \frac{1}{2} \sum_{i=1}^{n} \ln(\sigma_v^2 + \sigma_u^2 \sum_{t=1}^{T} g^2(EF \_it; \eta)) \]

\[ + \sum_{i=1}^{n} \ln[1 - \Phi(-\frac{\hat{u}_i}{\sigma_v})] + \frac{1}{2} \sum_{i=1}^{n} (\frac{\hat{u}_i}{\sigma_v})^2 - \frac{\sum_{i=1}^{n} \sum_{t=1}^{T} \hat{c}_i^2(\theta)}{2\sigma^2}. \] 

(2.5)

where \( \sigma_v^2 = \sigma_u^2 + \sigma^2 \sum_{t=1}^{T} g^2(EF \_it; \eta) \), \( \lambda = \sigma_u/\sigma_v \), \( \mu_u = -\frac{\sigma_u}{\sigma^2} \sum_{t=1}^{T} \epsilon_t \mu(EF \_it; \eta)/\sigma^2 \), \( \sigma_u^2 = \frac{\sigma_u^2}{\sigma_v^2} \), \( \sigma_v^2 = \frac{\sigma_u^2}{\sigma_v^2} \), \( \sigma^2 = \frac{\sigma_u^2}{\sigma_v^2} \), and \( \sigma^2 = \frac{\sigma_u^2}{\sigma_v^2} \). We estimate technical efficiency as \( \hat{TE} \_it = \frac{1}{\hat{Y} \_it - \hat{Q} \_it} \) and \( \hat{Q} \_it \)

\[ \hat{Y} \_it + \mu(EF \_it; \hat{\theta}) + \epsilon \_it, \] 

(2.6)

Again, since \( E(\hat{\epsilon} \_it | EF \_it) = 0 \), we estimate \( \hat{\delta}(EF \_it) = (\hat{\alpha}(EF \_it), \hat{\beta}(EF \_it)) \) with the standard smooth varying coefficient estimation. Specifically, \( \hat{\delta}(EF) = \hat{\alpha}_0 \), where \( \hat{\alpha}_0 = (\hat{\alpha}_{0,0}, \hat{\alpha}_{0,H}, \hat{\alpha}_{0,K}) \), \( \hat{\alpha}_1 = (\hat{\alpha}_{1,0}, \hat{\alpha}_{1,H}, \hat{\alpha}_{1,K}) \), \( \hat{\alpha}_2 = (\hat{\alpha}_{2,0}, \hat{\alpha}_{2,H}, \hat{\alpha}_{2,K}) \), such that

\[ \hat{\alpha} = \arg\min_{a} \sum_{i=1}^{n} \sum_{t=1}^{T} K \left( \frac{EF \_it - EF}{h} \right) (Y \_it - \hat{\bar{Q}} \_it) \] 

(2.7)

where \( \hat{\bar{Q}} \_it \) and \( \hat{Y} \_it = Y \_it + \mu(EF \_it; \hat{\theta}) \).

Finally, following Jondrow et al. (1982), we calculate observation specific technical inefficiencies for our semiparametric model. With proper modifications, we derive the conditional density of \( u \_it \) given \( \epsilon \_it \) and \( EF \_it \) as

\[ f(u \_it | \epsilon \_it, EF \_it) = [1 - \Phi(\epsilon \_it, \sigma(EF \_it), \sigma_v)]^{-1} \frac{1}{\sigma(EF \_it)} \exp[-\frac{1}{\sigma^2}(u \_it + \epsilon \_it, \sigma(EF \_it))^2] \]

where \( g \_it = g(EF \_it; \eta) \), and \( \sigma^2(EF \_it) = \sigma_v^2 + \sigma_u^2 g^2(EF \_it) \). We estimate technical efficiency as \( \hat{TE} \_it = e^{-M(u \_it | \epsilon \_it, EF \_it)} \), where \( M(u \_it | \epsilon \_it, EF \_it) = -\epsilon \_it, \sigma^2 g^2(EF \_it, \eta) \) if \( \epsilon \_it \leq 0 \) and \( M(u \_it | \epsilon \_it, EF \_it) = 0 \) if \( \epsilon \_it > 0 \).

In Equation (1), economic freedom shifts production frontier directly via \( \alpha(EF \_it) \) and indirectly through the output elasticities of input variables, \( \beta \_it(EF \_it), \beta \_it(EF \_it), \) and \( \beta \_it(EF \_it) \). Economic freedom level also affects technical efficiency through \( g(EF \_it; \eta) \) in \( M(u \_it | \epsilon \_it, EF \_it) \).
We compare our semiparametric model with three benchmark models. The first is a Cobb-Douglas production function with composite error which nests in Equation (2.1),

\[ Y_{it} = \alpha_0 + \alpha_1 EF_{it} + \alpha_2 EF_{it}^2 + \beta_1 \text{Ln}(H_{it}) + \beta_2 \text{Ln}(L_{it}) + \beta_3 \text{Ln}(K_{it}) + v_{it} - u_i e^{\eta_{EF_{it}}}, \] (2.8)

where we only allow economic freedom to shift the production frontier through the intercept term \( \alpha(EF_{it}) = \alpha_0 + \alpha_1 EF_{it} + \alpha_2 EF_{it}^2 \), a quadratic function to capture potential nonlinearity. All coefficients of \( X_{it} \) are assumed to be constants.

The second model is a translog model

\[ Y_{it} = \alpha_0 + \alpha_1 EF_{it} + \alpha_2 EF_{it}^2 + \beta_1 \text{Ln}(H_{it}) + \beta_2 \text{Ln}^2(H_{it}) + \beta_3 \text{Ln}(K_{it}) + \beta_4 \text{Ln}^2(K_{it}) + \beta_5 \text{EF}_{it} \text{Ln}(H_{it}) + \beta_6 \text{EF}_{it} \text{Ln}(K_{it}) + \beta_7 \text{Ln}(H_{it}) \text{Ln}(K_{it}) + v_{it} - u_i e^{\eta_{EF_{it}}}, \] (2.9)

Due to the presence of square and cross products of \( X_{it} \), it is not nested in Equation (2.1), but it imposes no priori restrictions on substitution possibilities among input variables.

In the third benchmark, we restrict the translog production model such that the coefficients of \( \text{Ln}^2(H_{it}), \text{Ln}^2(L_{it}), \text{Ln}^2(K_{it}), \text{Ln}(L_{it}) \text{Ln}(K_{it}), \text{Ln}(L_{it}) \text{Ln}(H_{it}), \text{Ln}(H_{it}) \text{Ln}(K_{it}) \) are jointly equal to zero, which makes it nested in Equation (2.1). After removing those terms, we estimate restricted translog model as,

\[ Y_{it} = \alpha_0 + \alpha_1 EF_{it} + \alpha_2 EF_{it}^2 + \beta_1 \text{Ln}(H_{it}) + \beta_2 \text{Ln}(L_{it}) + \beta_3 \text{Ln}(K_{it}) + \beta_4 \text{EF}_{it} \text{Ln}(H_{it}) + \beta_5 \text{EF}_{it} \text{Ln}(K_{it}) + v_{it} - u_i e^{\eta_{EF_{it}}}, \] (2.10)

### 2.3 Data Descriptions

We use the country-level data from the Penn World Table (PWT) and Economic Freedom of the World: 2015 Annual Report by Gwartney et al. (2015) to investigate the effects of economic freedom on a country's production frontier and technical efficiency. Our dataset consists of 658 observations for 94 countries from 1980 to 2010 observed at 5-year increments.\(^6\) A complete list of countries and their economic freedom as of 2010 are presented in Table 2.1. The EFW index is based on a 0 – 10 scale, with higher scores representing higher levels of economic freedom.

We employ the Penn World Table because it is the most comprehensive country-level aggregated data comprised of human capital per capita, labor, and physical capital. The latest version (8.1) of the Penn

\(^6\)We look at only five-year increments for two reasons. First, institutions change slowly and thus 5-year increments is standard in the economic freedom/growth literature. Second, the EFW index only reports at 5-year intervals from 1980 to 2000.
World Table provides output-side real GDP in million of 2005 US dollars for different countries over time, which facilitates comparison of economic productivity across a large number of countries. The combined data set enables us to estimate production frontiers and technical efficiency, and to investigate the effects of changing economic freedom on production frontiers.

Our dependent variable $Y$ is the logarithm of output-side real GDP in millions of 2005 US dollars. Independent variables include the logarithm of human capital per capita $\ln(H)$, the logarithm of labor force $\ln(L)$, and the logarithm of real capital stock $\ln(K)$, where $H$ is an index of human capital based on years of schooling and returns to education (see Psacharopoulos (1994) and Barro and Lee (2013)), labor force $L$ is in millions of persons participated in employment, and real capital stock $K$ is in millions of 2005 US dollars. Additional descriptions of variables from the Penn World Table can be obtained in Feenstra et al. (2015).

Economic freedom ($EF$) is our primary variable of interest in this paper and we treat it as the “environmental” variable in our estimation of production functions. Based on our hypothesis, the effects of economic freedom on output are double-sided. Increases in economic freedom can reduce transaction costs for productive activities, make foreign capital more accessible and domestic capital more productive, and improve the return on education, all of which boost productivity. However, on the other hand, increases in economic freedom also relax governmental control of the population, which increases the migration of immigrants to countries with higher levels of economic freedom. While this is likely to increase the productivity of destination countries, it will lower output elasticity in the origin country.

We use the chain-linked EFW index ($EFW$) to ensure comparability across time. $EFW$ measures the degree of economics freedom in five major areas: (1) size of government; (2) legal system and security of property rights; (3) access to sound money; (4) freedom to trade internationally; and (5) regulation. Each of these components are based on several variables. The rating of size of government, for instance, is based on four separate components, such as government consumption as a percentage of total consumption. Each component is put on a 0 to 10 scale and then aggregated up to an overall score for the entire country, which also varies from 0 to 10. The detailed $EFW$ index structure can be obtained from Gwartney et al. (2015). Gwartney and Lawson (2003) provides an overview of the history and creation of the index. The summary statistics for all the variables are presented in Table 2.2.

### 2.4 Estimation Results

The estimation results for our semiparametric smooth coefficient model are summarized in Table 1.8, providing the mean values and 10th, 50th (median), and 90th percentile of our smooth coefficients of $\hat{\delta}(EF) =$
Chapter 2. Does Economic Freedom Affect The Production Frontier?

\[ \hat{\alpha}(EF), \hat{\beta}_H(EF), \hat{\beta}_L(EF), \hat{\beta}_K(EF) \] \', as well as the parameter estimate \[ \hat{\theta} = (\hat{\sigma}_u^2, \hat{\sigma}_v^2, \hat{\eta})' \]. Compared with benchmark models presented in Table 2.4, our semiparametric estimates give much higher \[ \hat{\sigma}_u^2 \] and \[ \hat{\sigma}_v^2 \].

Compared with benchmark models presented in Table 2.4, our semiparametric estimates give much higher \[ \hat{\sigma}_u^2 \] and \[ \hat{\sigma}_v^2 \], .7274 and .0489, respectively. The Cobb-Douglas production frontier, translog frontier, and restricted translog frontier give relative lower estimates of \[ (\hat{\sigma}_u^2, \hat{\sigma}_v^2) = (\hat{\sigma}_u^2, \hat{\sigma}_v^2) = (.2086, .0341), (\hat{\sigma}_u^2, \hat{\sigma}_v^2) = (.1710, .0290), \] and \[ (\hat{\sigma}_u^2, \hat{\sigma}_v^2) = (.6488, .0324), \] respectively. All estimates suggest a relatively small magnitude for the random noise relative to that of the efficiency term. With an estimated \[ \hat{\eta} \] of \(-.0674\), our semiparametric model implies that increased economic freedom will increase the technical efficiency \( (TE_t) \). The restrictive nature of the Cobb-Douglas and translog, however, deliver an estimated \[ \hat{\eta} \] of \(.0902 \) and \(.0946 \), respectively, implying that increased economic freedom actually reduces efficiency, which is counter-intuitive. The restricted translog model, on the other hand, gives an estimated \[ \hat{\eta} \] of \(-.0156 \), confirming our semiparametric estimate.

The restricted translog production function assumes that the coefficients of \( \ln^2(H_{it}), \ln^2(L_{it}), \ln(K_{it}), \ln(L_{it})\ln(K_{it}), \) and \( \ln(H_{it})\ln(K_{it}) \) are jointly equal to zero. To check its validity empirically, we perform a likelihood ratio test for the null of \[ \beta_H^2 = \beta_L^2 = \beta_K^2 = \beta_{HL} = \beta_{HK} = \beta_{LK} = 0 \] which gives us a test statistic of \(.0975 \), smaller than the critical value of \(12.592 \) at \(5\% \) significance level. Since we cannot reject the null hypothesis and the restricted translog model appears to be a better fit compared with translog model, we focus on Cobb-Douglas and restricted translog models as parametric counterparts in the following analysis.

Since our semiparametric estimators are smooth functions of the environmental variable \( EF \), we plot our coefficients \[ (\hat{\alpha}(EF), \hat{\beta}_H(EF), \hat{\beta}_L(EF), \hat{\beta}_K(EF))' \] and their \(95\% \) confidence bounds, which are based on the asymptotic results in Yao et al. (2016), against economic freedom in Figure 2.1. For comparison purposes, we superimpose the parametric estimates from the Cobb-Douglas and restricted translog models. As expected, our smooth coefficients are mostly positive and highly nonlinear across the span of economic freedom, and they do exhibit different patterns.

The left-upper panel in Figure 2.1 presents the direct effects of economic freedom \( \hat{\alpha}(EF) \) against its parametric counterparts. Once we allow the coefficients to vary with economic freedom, the direct effects of economic freedom on production are much smaller than those from parametric models. The right-upper panel in Figure 2.1 describes the output elasticity of human capital (or marginal product of human capital) against economic freedom. It shows a decreasing pattern in general but also suggests improved productivities from human capital when economic freedom is below \( 4.4 \) (Venezuela in 2010) and above \( 7.5 \) (approximately Japan in 2010). The \(95\% \) confidence bounds of semiparametric estimates contain most of its parametric counterparts, suggesting parametric models yield fairly reasonable estimates for the output elasticity of human capital.
The left-lower panel in Figure 2.1 plots the output elasticity of labor (or marginal product of labor). It shares a similar pattern with that of human capital, increasing first when economic freedom is low then decreasing in a wide range until economic freedom reaches the threshold around 7.75, however, our semiparametric estimates are much lower than the parametric estimates. The output elasticity of capital in the right-lower panel exhibits an increasing pattern in general until economic freedom reaches the threshold of 7.75 (Ireland in 2010). Our results suggests higher output elasticity of capital under semiparametric specification than parametric counterparts.

Discrepancy between semiparametric and parametric estimates implies that the parametric approaches, which are very much likely to suffer from misspecification, overestimate the direct effect of economic freedom, output elasticity of labor and underestimate output elasticity of capital, but deliver reasonable estimates of output elasticity of human capital.

Given the pattern exhibited in Figure 1 and the fact that $Ln(K)$’s magnitude is larger than other inputs, we expect that economic freedom shifts the production frontier upward. For illustration, let’s consider the case of China. In 2010, China had a chain-linked EF score of 6.07. Our semiparametric model suggests that if China has the economic freedom level of the United States (7.76), its production frontier would be shifted upward by $4,047,534$ millions of 2005 US dollar or 21.39% higher than its estimated 2010 production frontier, ceteris paribus. The Cobb-Douglas and restricted translog models suggest that the same increase in EF will have a larger impact, with an upward shift of $4,566,835$ millions of 2005 US dollar (or 29.78% higher than the estimated 2010 production frontier), and an upward shift of $5,642,097$ millions of 2005 US dollar (or 38.19% higher), respectively.

Figure 2.2 presents kernel density estimates of composite error and technical efficiency for all three models. The left panel indicates that composite errors for all three models cluster around $-0.5$ which suggests majority of observations are not fully technically efficient on average ($E(\epsilon_{it}|EF_{it}) < 0$) and operates below production frontier. Moreover, the kernel density of composite error for semiparametric estimates is taller and more tightly centered around $-0.5$, suggesting smaller composite error in absolute values and higher technical efficiency on average under semiparametric estimates than those of parametric counterparts.

The right panel of Figure 2.2 presents density estimates of technical efficiency for semiparametric and parametric estimates. The results are consistent with what is suggested in the left panel. Compared with parametric counterparts, the density of semiparametric technical efficiency is more tightly centered at a higher level, suggesting higher efficiencies under the semiparametric approach. The average and median technical efficiency are .7146 and .7197 for the semiparametric model, .5074 and .4866 for Cobb-Douglas model, and .5511 and .5494 for translog model. The results suggest that technical efficiencies are likely
to be underestimated in parametric models. Our results are consistent with Adkins et al. (2002) and Klein and Luu (2003) in the sense that majority of observations are technically inefficient. However, our estimates of technical efficiency are smaller in absolute values. A direct comparison between our and their estimates may not be appropriate since they incorporate many other variables such as political rights and civil liberty in the estimation of technical efficiency. However, we conjecture that their estimates can be misleading, since they assume constant output elasticities which can either overestimate or underestimate the impact of economic freedom on the elasticities as demonstrated in Figure 1.

Following Li et al. (2002) and Li and Racine (2010), we perform a modified model specification test (See Yao et al. (2016)) for the parametric models against our semiparametric model. The test statistics $\hat{T}_n$ for the null of Cobb-Douglas and translog model are 362.75 and 153.6, respectively, and the bootstrapped $p-$value are zero under both scenarios, strongly suggesting parametric models are misspecified.

2.5 Conclusion

In this paper, we apply a multi-step semiparametric smooth coefficient stochastic production frontier estimator proposed by Yao et al. (2016) to investigate the effects of economic freedom on production frontier and technical efficiency. We contrast the semiparametric estimate with Cobb-Douglas and translog. Allowing the output elasticities and technical efficiency to depend on the environmental variable, economic freedom, we observe significant variation on output elasticities. A model specification test indicates semiparametric smooth coefficient stochastic production frontier is a better fit than both Cobb-Douglas and translog production frontiers. Like the previous literature we find that economic freedom shifts the semiparametric stochastic production frontier upward and reduces technical inefficiency.

Our results add to the literature on economic freedom and growth in two ways. First, our results highlight the importance of semiparametric approaches as we find that parametric approaches to estimating the marginal productivity of inputs to be restrictive. Parametric approaches overestimate direct effect of economic freedom and output elasticity of labor, and underestimate the output elasticity of capital. Second, we find that the output elasticities of labor, human capital, and physical capital vary with the level of economic freedom. In this way our results are similar to that of Hall et al. (2010), but more precise. Our results suggest that the output elasticities are mostly positive. Interestingly, increased economic freedom generally lowers the output elasticities or marginal products of human capital and labor but leads to improvements in the marginal product of capital.
Table 2.1: List of Countries According to Economic Freedom (as of 2010)

<table>
<thead>
<tr>
<th>Country</th>
<th>EF</th>
<th>Country</th>
<th>EF</th>
<th>Country</th>
<th>EF</th>
<th>Country</th>
<th>EF</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 &lt; $EF \leq 6.5$</td>
<td></td>
<td>6.5 &lt; $EF \leq 7.5$</td>
<td></td>
<td>7.5 &lt; $EF \leq 10$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Argentina</td>
<td>5.86</td>
<td>Bangladesh</td>
<td>6.52</td>
<td>Philippines</td>
<td>7.09</td>
<td>Australia</td>
<td>8.10</td>
</tr>
<tr>
<td>Belize</td>
<td>6.45</td>
<td>Barbados</td>
<td>6.63</td>
<td>Portugal</td>
<td>7.06</td>
<td>Austria</td>
<td>7.53</td>
</tr>
<tr>
<td>Benin</td>
<td>5.84</td>
<td>Belgium</td>
<td>7.47</td>
<td>Sierra Leone</td>
<td>6.91</td>
<td>Bahrain</td>
<td>7.69</td>
</tr>
<tr>
<td>Bolivia</td>
<td>6.39</td>
<td>Botswana</td>
<td>7.22</td>
<td>South Africa</td>
<td>6.87</td>
<td>Canada</td>
<td>8.05</td>
</tr>
<tr>
<td>Burundi</td>
<td>5.02</td>
<td>Brazil</td>
<td>6.55</td>
<td>Spain</td>
<td>7.26</td>
<td>Chile</td>
<td>7.94</td>
</tr>
<tr>
<td>China</td>
<td>6.07</td>
<td>Cameroon</td>
<td>6.56</td>
<td>Tanzania</td>
<td>6.54</td>
<td>Cyprus</td>
<td>7.60</td>
</tr>
<tr>
<td>Colombia</td>
<td>6.33</td>
<td>Costa Rica</td>
<td>7.41</td>
<td>Thailand</td>
<td>6.66</td>
<td>Denmark</td>
<td>7.75</td>
</tr>
<tr>
<td>Congo, Dem.</td>
<td>5.61</td>
<td>Dominican Rep.</td>
<td>7.06</td>
<td>Trinidad &amp; Tob.</td>
<td>6.99</td>
<td>Finland</td>
<td>7.73</td>
</tr>
<tr>
<td>Congo, Rep.</td>
<td>4.92</td>
<td>Egypt</td>
<td>6.79</td>
<td>Turkey</td>
<td>6.52</td>
<td>Germany</td>
<td>7.53</td>
</tr>
<tr>
<td>Cote d’Ivoire</td>
<td>5.91</td>
<td>El Salvador</td>
<td>7.43</td>
<td>Uruguay</td>
<td>7.29</td>
<td>Hong Kong</td>
<td>8.86</td>
</tr>
<tr>
<td>Ecuador</td>
<td>5.70</td>
<td>France</td>
<td>7.43</td>
<td></td>
<td></td>
<td>Ireland</td>
<td>7.75</td>
</tr>
<tr>
<td>Fiji</td>
<td>6.36</td>
<td>Ghana</td>
<td>6.67</td>
<td></td>
<td></td>
<td>Israel</td>
<td>7.60</td>
</tr>
<tr>
<td>Gabon</td>
<td>5.63</td>
<td>Greece</td>
<td>6.75</td>
<td></td>
<td></td>
<td>Japan</td>
<td>7.51</td>
</tr>
<tr>
<td>Iceland</td>
<td>6.41</td>
<td>Guatemala</td>
<td>7.24</td>
<td></td>
<td></td>
<td>Jordan</td>
<td>7.75</td>
</tr>
<tr>
<td>Iran</td>
<td>6.41</td>
<td>Honduras</td>
<td>7.04</td>
<td></td>
<td></td>
<td>Luxembourg</td>
<td>7.58</td>
</tr>
<tr>
<td>Malawi</td>
<td>5.97</td>
<td>Hungary</td>
<td>7.32</td>
<td></td>
<td></td>
<td>Malta</td>
<td>7.61</td>
</tr>
<tr>
<td>Mali</td>
<td>5.93</td>
<td>India</td>
<td>6.59</td>
<td></td>
<td></td>
<td>Mauritius</td>
<td>7.92</td>
</tr>
<tr>
<td>Morocco</td>
<td>6.45</td>
<td>Indonesia</td>
<td>7.05</td>
<td></td>
<td></td>
<td>Netherlands</td>
<td>7.58</td>
</tr>
<tr>
<td>Nepal</td>
<td>5.92</td>
<td>Italy</td>
<td>7.11</td>
<td></td>
<td></td>
<td>New Zealand</td>
<td>8.10</td>
</tr>
<tr>
<td>Niger</td>
<td>5.92</td>
<td>Jamaica</td>
<td>7.12</td>
<td></td>
<td></td>
<td>Peru</td>
<td>7.51</td>
</tr>
<tr>
<td>Pakistan</td>
<td>5.98</td>
<td>Kenya</td>
<td>7.07</td>
<td></td>
<td></td>
<td>Singapore</td>
<td>8.53</td>
</tr>
<tr>
<td>Senegal</td>
<td>5.78</td>
<td>Korea, South</td>
<td>7.28</td>
<td></td>
<td></td>
<td>Sweden</td>
<td>7.61</td>
</tr>
<tr>
<td>Sri Lanka</td>
<td>6.25</td>
<td>Kuwait</td>
<td>7.40</td>
<td></td>
<td></td>
<td>Switzerland</td>
<td>8.23</td>
</tr>
<tr>
<td>Syria</td>
<td>5.72</td>
<td>Malaysia</td>
<td>7.00</td>
<td></td>
<td></td>
<td>Taiwan</td>
<td>7.74</td>
</tr>
<tr>
<td>Togo</td>
<td>5.61</td>
<td>Mexico</td>
<td>6.69</td>
<td></td>
<td></td>
<td>Uganda</td>
<td>7.56</td>
</tr>
<tr>
<td>Tunisia</td>
<td>6.06</td>
<td>Norway</td>
<td>7.43</td>
<td></td>
<td></td>
<td>United Kingdom</td>
<td>7.91</td>
</tr>
<tr>
<td>Venezuela</td>
<td>3.84</td>
<td>Panama</td>
<td>7.25</td>
<td></td>
<td></td>
<td>United States</td>
<td>7.76</td>
</tr>
<tr>
<td>Zimbabwe</td>
<td>4.38</td>
<td>Paraguay</td>
<td>6.62</td>
<td></td>
<td></td>
<td>Zambia</td>
<td>7.68</td>
</tr>
</tbody>
</table>
### Table 2.2: Summary Statistics

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Variable Description</th>
<th>Mean</th>
<th>Std.</th>
<th>Min.</th>
<th>Max.</th>
<th>Bandwidth</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\ln(Y)$</td>
<td>Log of Output-Side Real GDP</td>
<td>11.1473</td>
<td>1.9426</td>
<td>6.2424</td>
<td>16.3800</td>
<td>–</td>
</tr>
<tr>
<td>$\ln(H)$</td>
<td>Log of Human Capital</td>
<td>.8102</td>
<td>.2662</td>
<td>.0827</td>
<td>1.2861</td>
<td>–</td>
</tr>
<tr>
<td>$\ln(L)$</td>
<td>Log of Labor Force</td>
<td>1.4582</td>
<td>1.7037</td>
<td>−3.2114</td>
<td>6.6611</td>
<td>–</td>
</tr>
<tr>
<td>$\ln(K)$</td>
<td>Log of Capital Stock</td>
<td>11.9540</td>
<td>2.1482</td>
<td>6.3503</td>
<td>17.5224</td>
<td>–</td>
</tr>
<tr>
<td>$EF$</td>
<td>Economics Freedom</td>
<td>6.2450</td>
<td>1.3167</td>
<td>2.6128</td>
<td>9.1510</td>
<td>.5307</td>
</tr>
</tbody>
</table>

### Table 2.3: Estimation of the Semiparametric Smooth Coefficient Stochastic Frontier

<table>
<thead>
<tr>
<th>Smooth Coefficient</th>
<th>MLE Estimates</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\sigma}_u^2$</td>
<td>.7274</td>
<td>.3366</td>
</tr>
<tr>
<td>$\hat{\sigma}_v^2$</td>
<td>.0489</td>
<td>.0058</td>
</tr>
<tr>
<td>$\hat{\eta}$</td>
<td>−.0674</td>
<td>.0012</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Smooth Coefficient</th>
<th>Mean</th>
<th>10th Percentile</th>
<th>Median</th>
<th>90th Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\alpha}(EF)$</td>
<td>2.3659</td>
<td>1.2929</td>
<td>2.3263</td>
<td>3.6821</td>
</tr>
<tr>
<td>$\hat{\beta}_H(EF)$</td>
<td>.3232</td>
<td>−.0334</td>
<td>.2769</td>
<td>.8246</td>
</tr>
<tr>
<td>$\hat{\beta}_L(EF)$</td>
<td>.2438</td>
<td>.1517</td>
<td>.2197</td>
<td>.3899</td>
</tr>
<tr>
<td>$\hat{\beta}_K(EF)$</td>
<td>.7189</td>
<td>.5553</td>
<td>.7274</td>
<td>.8361</td>
</tr>
</tbody>
</table>
Table 2.4: Maximum Likelihood Estimation of Parametric Benchmark Models

<table>
<thead>
<tr>
<th>Variables</th>
<th>Cobb-Douglas Model</th>
<th>Translog Model</th>
<th>Restricted Translog Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>5.2642</td>
<td>.0000</td>
<td>10.5165</td>
</tr>
<tr>
<td>$EF$</td>
<td>−.1651</td>
<td>.0101</td>
<td>.0306</td>
</tr>
<tr>
<td>$EF^2$</td>
<td>.0231</td>
<td>.0000</td>
<td>.0096</td>
</tr>
<tr>
<td>$\ln(H)$</td>
<td>.2535</td>
<td>.0127</td>
<td>2.1816</td>
</tr>
<tr>
<td>$\ln(L)$</td>
<td>.3985</td>
<td>.0000</td>
<td>1.4784</td>
</tr>
<tr>
<td>$\ln(K)$</td>
<td>.4879</td>
<td>.0000</td>
<td>−.8103</td>
</tr>
<tr>
<td>$\ln^2(H)$</td>
<td>−</td>
<td>−</td>
<td>1.2049</td>
</tr>
<tr>
<td>$\ln^2(L)$</td>
<td>−</td>
<td>−</td>
<td>.0651</td>
</tr>
<tr>
<td>$\ln^2(K)$</td>
<td>−</td>
<td>−</td>
<td>.0831</td>
</tr>
<tr>
<td>$EFLn(H)$</td>
<td>−</td>
<td>−</td>
<td>.0876</td>
</tr>
<tr>
<td>$EFLn(L)$</td>
<td>−</td>
<td>−</td>
<td>.0269</td>
</tr>
<tr>
<td>$EFLn(K)$</td>
<td>−</td>
<td>−</td>
<td>−.0138</td>
</tr>
<tr>
<td>$\ln(H)\ln(L)$</td>
<td>−</td>
<td>−</td>
<td>.4401</td>
</tr>
<tr>
<td>$\ln(H)\ln(K)$</td>
<td>−</td>
<td>−</td>
<td>−.4202</td>
</tr>
<tr>
<td>$\ln(L)\ln(K)$</td>
<td>−</td>
<td>−</td>
<td>−.1522</td>
</tr>
<tr>
<td>$\hat{\sigma}^2_u$</td>
<td>.2086</td>
<td>.0005</td>
<td>.1710</td>
</tr>
<tr>
<td>$\hat{\sigma}^2_v$</td>
<td>.0341</td>
<td>.0000</td>
<td>.0290</td>
</tr>
<tr>
<td>$\hat{\eta}$</td>
<td>.0002</td>
<td>.0000</td>
<td>.0946</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>546.406</td>
<td>598.967</td>
<td>570.474</td>
</tr>
<tr>
<td>Observations</td>
<td>658</td>
<td>658</td>
<td>658</td>
</tr>
</tbody>
</table>
Chapter 2. Does Economic Freedom Affect The Production Frontier?

Figure 2.1: Plots of Smooth Coefficient Frontiers

Figure 2.2: Kernel Density of Composite Error and Technical Efficiency
Bibliography


Chapter 3

Conditional Factor Model: A Semiparametric Approach

3.1 Introduction

Previous theoretical research has argued that factor loadings in asset pricing models are not constant; that is, alphas and betas increase during times when risk premia are high. Specifically, several theoretical models of risk suggest that betas on value stocks should be time-varying and should be highest during time periods in which marginal utility is high. While much of the previous research utilized rolling windows to check for the time variation in betas, Ang and Kristensen (2012) and Li and Yang (2011) use nonparametric techniques and to test whether the betas are time-varying in two of the post prominent factor pricing models (CAPM and the Fama French Three Factor Model). Both papers report that the factor loadings for the CAPM and the Fama-French (henceforth FF) three factor model are time varying. However, in both papers, after accounting for time varying alphas (abnormal returns) and betas, neither model is able to many asset pricing anomalies (i.e. rejection that the alphas are equal to zero). The FF three factor model captures the relationship between the average return and book-to-market ($B/M$) and the relationship between the average return and size. As such, one possibility for the pricing anomalies was provided by Novy-Marx (2013) in which it is argued that that the traditional three factor model is incomplete because it fails to account for the relationship between average returns that is related to profitability and investment. Thus, Fama and French (2015) recently included two additional factors to their seminal three factor model (RMW and CMA); where RMW is calculated as the difference between the returns of equity portfolios of robust and weak profitability and CMA is the difference between returns of equity portfolios of low and high investment firms.
Our primary aim in this paper is to build upon the work of Ang and Kristensen (2012) and Li and Yang (2011) by using semiparametric techniques to allow the alphas and betas in the five factor FF model to be variant. However, we allow them in the FF five factor model to be functions of the real interest rate. That is, we acknowledge that the factor loadings are time-varying; however, we believe that the theoretical literature (such as Lettau and Ludvigson (2011)) suggests that the factor loadings are changing through time due to changes in the risk premium or changes in marginal utility. For example, recent debates in the U.S regarding the decline in the equilibrium real interest rate due to secular stagnation over the past three decades would suggest a change in the marginal utility of consumption. As such, we estimate the FF five factor model using smooth varying coefficient model pioneered by Li et al. (2002) where the alphas and betas are functions of the real interest rate.

The remainder of this article is organized as following. Section 2 introduces our methodology and model specifications. Section 3 describes the nature of variables and data sources. Section 4 discusses empirical results, illustrating the smooth varying coefficients of conditional Fama–French five factor model and comparing results with their OLS counterparts. Section 5 concludes.

3.2 Methodology and Model Specification

Ang and Kristensen (2012) and Li and Yang (2011) develop conditional factor models contemporaneously using nonparametric method and drive time-varying conditional alphas and betas. However, only focusing on time-varying conditional alphas and betas, they ignored the effects many other fundamental macroeconomic variables, such real interest rate, on conditional factor loadings. Ferreira et al. (2011) also employ time-varying conditional factor models, however, they still ignore the effects of macroeconomic variables. Consider the following semiparametric smooth varying coefficient model, for portfolio $i$ and time $t = 1, \cdots, T$,

$$R_{it} - R_{ft} = \alpha_i(r_t) + \beta_i(r_t)X_t + \epsilon_{it},$$

where $R_{it} - R_{ft}$ is the excess return for portfolio $i$ at time $t$, $X_t = (MKT_t, SMB_t, HML_t, RMW_t, CMA_t)'$ represents the market excess return $(R_{Mt} - R_{ft})$ and returns for the mimicking portfolios for the size, book-to-market, profitability and investment factors. $\beta_i(r_t) = (\beta_M(r_t), \beta_S(r_t), \beta_H(r_t), \beta_R(r_t), \beta_C(r_t))'$ and $\delta_i(r_t) = (\alpha_i(r_t), \beta_i(r_t))'$ is a vector of unknown smooth functions of exogenous environmental variable, the real interest rate. Smooth coefficient function increases the flexibility that a traditional Fama–French five factor model cannot afford. $\epsilon_{it}$ is a random noise and $E(\epsilon_{it}|r_t) = 0$. we apply standard smooth varying coefficient estimation (see Li et al. (2002)) on Equation (1.5) to obtain consistent estimator $\hat{\delta}(r) = \hat{b}_0$, where $\hat{b}_0 = (\hat{b}_{00}, \hat{b}_{0M}, \hat{b}_{0S}, \hat{b}_{0H}, \hat{b}_{0R}, \hat{b}_{0C})'$, $\hat{b}_1 = (\hat{b}_{10}, \hat{b}_{1M}, \hat{b}_{1S}, \hat{b}_{1H}, \hat{b}_{1R}, \hat{b}_{1C})'$,
\( \hat{b} = (\hat{b}_0, \hat{b}_1)' \), such that
\[
\hat{b} = \arg\min_b \sum_{t=1}^{T} K \left( \frac{r_t - r}{h} \right) \left( y_{it} - \hat{W}_t' b \right)^2,
\]
(3.2)
where \( \hat{W}_t = \begin{bmatrix} (1, X_t')' \\ (1, X_t')' \otimes (r_t - r) \end{bmatrix} \), \( \otimes \) denotes the Kronecker product, and \( y_{it} = R_{it} - R_{ft} \). \( K(\cdot) \) represents a weighting schedule which assigns different weights to observations based on their distances to the point of estimation. In the empirical work, We employ a Gaussian kernel function \( K(u) = e^{-u^2/2}/\sqrt{2\pi} \) to allocate larger weights to observations that close to the point of estimation and smaller weights to observations that far away from the point of estimation. To select the optimal window, we follow standard automatic rule-of-thumb bandwidth selection method by Ruppert et al. (1995) which provides better results than cross-validation methods in both asymptotics and practical performance.

Moreover, following the practice in Li et al. (2002), the variances of smooth coefficients for portfolio \( i \) can be consistently estimated as \( \hat{\Omega}_i = \hat{Q}_i^{-1} \hat{V}_i \hat{Q}_i^{-1} \), where \( \hat{Q} = (nh)^{-1} \sum_{t=1}^{T} X_t X_t' K_{tr} \) and \( K_{tr} = K \left( \frac{r_t - r}{h} \right) \), \( \hat{V}_i = (nh)^{-1} \sum_{t=1}^{T} X_t X_t' \hat{\epsilon}_t^2 K_{tr}^2 \), and \( \hat{\epsilon}_t = y_{it} - X_t' \hat{\delta}(r_t) \).

We compare our semiparametric smooth coefficient results with traditional Fama–French five factor model to explores their difference and similarities. We apply OLS estimation on equation (3.3) to derive traditional Fama–French five factor results for portfolio \( i \),
\[
R_{it} - R_{ft} = \alpha_i + \beta_{iM} MKT_t + \beta_{iS} SML_t + \beta_{iH} HML_t + \beta_{iR} RMW_t + \beta_{iC} CMA_t + \epsilon_{it},
\]
(3.3)

### 3.3 Data Description

We use monthly returns on 18 portfolios (9 equally-weighted portfolios and 9 value-weighted portfolios) formed on size, book-to-market value \((B/M)\), and operating profitability \((OP)\) to investigate the effects of real interest rate on abnormal return \((\alpha)\) and conditional factor loadings. Our dataset consists of 636 monthly observations for 18 portfolios from July, 1963 to June, 2016. All the portfolio data and factors data are retrieved from Fama-French library (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html). The real interest rate is obtained from Fisher equation and we approximate real interest as \( r_t \approx i_t - \pi_t \). We use 10 – year treasury bond rate to approximate nominal interest rate and CPI as inflation rate. Monthly data for 10 – year treasury bond rate and CPI are retrieved from Federal Reserve St. Louis website (https://fred.stlouisfed.org/). Smooth coefficient model requires stationarity from all variables as prerequisite and we performed ADF test with Akalike info criterion and auto lag section. The summary statistics and ADF test statistics for all variables are presented in Table - 3.1. All the variables are stationary at 1% significance level except environmental variable real interest rate.
which is stationary at 10% significance level with a $P$-value of .0539.

We employ the univariate sorts data from Fama–French library because we can explore the effects of real interest rate on specific types of portfolio. For example, we can probe how abnormal return of a value portfolio (high $B/M$) changes when real interest rate changes. A long time span of data also improves estimation efficiency and provides a more accurate results for long-run factor loadings in OLS estimation. The combined dataset enables us to estimate the smooth coefficient Fama–French model and investigate the effects of real interest rate on abnormal returns and factor loadings.

Our dependent variable $R_i - R_f$ is the excess return of portfolio $i$ in monthly frequency. Independent variables are $MKT$, the return from the value weighted market portfolio minus the risk free rate, $SMB$, the difference between the return of a diversified small size stocks portfolio and the return of a diversified large size stock portfolio, $HML$, the return of a diversified high $B/M$ stocks portfolio minus the return of a diversified low $B/M$ stocks portfolio, $RMW$, the return of a diversified high profitability firms portfolio minus the return of a diversified low profitability firms portfolio, and $CMA$, the difference between the return of a diversified low investment firms portfolio and the return of a diversified high investment firms portfolio.

Real interest rate is the variable of interest in this paper and it serves as the exogenous environmental variable in our estimation of conditional Fama–French factor model. Traditional conditional factor models employ a time schedule as environmental variable and derives time-varying factor loadings. This thread of literature implicitly investigates the effects of time on the asset pricing model and assumes time alone contains all the necessary information to price appropriately. However, a time schedule cannot simply capture all the determine factor in asset pricing. Instead, in a very diversified portfolio when risks from individual firms are no longer major concerns, we believe that macroeconomic variables such as real interest rate are the actual driving forces of asset pricing models.

### 3.4 Empirical Results

The estimation results of our conditional Fama-French five factor model for equally weighted portfolios and value weighted portfolios are summarized in Table - 3.2 and Table - 3.3, providing the 10th, 50th (median), and 90th percentile of our conditional factor loadings of $\hat{\delta}(r) = (\hat{\alpha}(r), \hat{\beta}_M(r), \hat{\beta}_S(r), \hat{\beta}_H(r), \hat{\beta}_R(r), \hat{\beta}_C(r))'$. The OLS results from traditional Fama-French model are presented in Table - 3.4 and Table - 3.5.

Estimation results from conditional factor model suggests very different investment strategies to generate abnormal return from its OLS counterparts. Equally weighted portfolios for instance, traditional
Fama–French model suggest we cannot generate large abnormal returns in long run except portfolio formed on high book-to-market value firms and portfolio formed with low profitability firms with annualized abnormal returns of 4.33% and 2.57%, respectively. No significant abnormal returns are observed in other equally weighted portfolios. However, results from conditional factor model indicate we can achieve higher abnormal returns for equally weighted portfolios (except portfolio formed with high operating profitability firms) at different real interest rate environment. For example, portfolio formed with high book-to-market firms can generate an annualized abnormal return of 5.47% (90th percentile) and portfolio formed with low profitability firms can generate an annualized abnormal return of 3.49% (90th percentile), both of them are higher than their OLS counter parts. Also, for portfolios which cannot generate abnormal returns in traditional factor models, our conditional factor model suggest abnormal returns can be achieved at specific real interest rate environment. Portfolios formed on small size firms and low book-to-market firms for example, conditional factor model suggests annualized abnormal returns of 2.45% and 2.36% at 90th percentile while traditional factor suggests abnormal returns cannot be achieved. In addition, traditional factor model does not indicate any significant negative abnormal returns (when market average performs better than individual portfolios) exist. However, conditional factor model suggests under certain real interest environment, market average outperforms portfolio formed with medium size firms, portfolio formed large size firms, and portfolio formed with low book-to-market firms by 1.98%, 1.34%, and 4.38%, respectively.

For value weighted portfolios, traditional factor model suggests positive abnormal returns cannot be achieved in long run and indicates an annualized negative abnormal return of 1.36% for portfolio formed with medium book-to-market firms. On the contrary, conditional factor model indicates positive abnormal returns can be achieved for portfolios formed on low book-to-market firms, on low operating profitability firms, and on high operating profitability firms with an annualized abnormal returns of 1.09%, 1.02%, and 1.26% at their 90th percentile. Conditional factor model also suggest significant negative abnormal returns for portfolio formed with medium book-to-market firms and portfolio formed with medium profitability firms.

Conditional factor model also indicates real interest rate varying coefficients for factor loadings while traditional Fama–French model suggest constant factor loadings. Take equally-weighed size portfolios for example, factor loadings for \( MKT \), \( HML \), \( RMW \), and \( CMA \) all exhibit different degrees of volatility against real interest except \( SMB \). Figure - 3.1 presents conditional factor loadings for equally-weighed size portfolios (small, medium, large) with respect to real interest rate. The largest variations for factor loadings can be observed in \( HML \), \( RMW \), and \( CMA \). Conditional factor loadings of \( HML \) for equally weighted small portfolio range from -.10 to .25 when real interest rates changes. It shows a lowest \( HML \).
factor loading when real interest rate is approximately -6% and shows the highest $HML$ loading when real interest rate is around 2%. Similarly, conditional factor loadings of $RMW$ for medium size portfolio range from -.4 to .15 when real interest rates changes. It also exhibits the lowest $RMW$ factor loading when real interest rate is approximately -6% and shows the highest $RMW$ loading when real interest rate is around 2%.

3.5 Conclusion

In this Chapter, we apply a smooth varying coefficient model proposed by Li et al. (2002) to investigate the effects of real interest rate on Fama–French five factor model. We contrast the conditional factor model with it’s traditional setup. Allowing the abnormal returns and factor loadings to depend on the real interest rate, we observe significant variation on abnormal returns and factor loadings in all the portfolios investigated when real interest rates changes. Our results add to the literature on factor pricing model in two ways. First, our results highlight the importance of the application of semiparametric methods in the estimation of conditional factor models due to significant variations found and restrictive nature of parametric method. In this way our results are similar to that of Ang and Kristensen (2012) and Li and Yang (2011), but more precise. Second, we find that the abnormal returns and factor loadings vary with the real interest instead of time. Our results indicates we can generate abnormal returns from different portfolios when real interest rate changes.
3.6 Appendix: Tables and Figures

Table 3.1: Summary Statistics

Panel A: factors

<table>
<thead>
<tr>
<th>Factor</th>
<th>Mean</th>
<th>S.T.D</th>
<th>ADF – Statistic</th>
<th>Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td>MKT</td>
<td>.5004</td>
<td>4.438</td>
<td>-23.39**</td>
<td>1</td>
</tr>
<tr>
<td>SMB</td>
<td>.2541</td>
<td>3.058</td>
<td>-23.56**</td>
<td>.2809</td>
</tr>
<tr>
<td>HML</td>
<td>.3429</td>
<td>2.850</td>
<td>-21.57**</td>
<td>-.2975</td>
</tr>
<tr>
<td>RMW</td>
<td>.2483</td>
<td>2.119</td>
<td>-21.11**</td>
<td>-.2052</td>
</tr>
<tr>
<td>CMA</td>
<td>.3093</td>
<td>2.011</td>
<td>-22.15**</td>
<td>-.3878</td>
</tr>
</tbody>
</table>

Panel B: Environmental Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>S.T.D</th>
<th>ADF – Statistic</th>
<th>Bandwidth Selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>i (10-year Treasury Bond)</td>
<td>6.428</td>
<td>2.833</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>π (CPI)</td>
<td>3.998</td>
<td>2.894</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>r (Real Interest Rate)</td>
<td>2.430</td>
<td>2.402</td>
<td>-2.856*</td>
<td>Ruppert et al. (1995)</td>
</tr>
</tbody>
</table>

Panel C: Portfolios

<table>
<thead>
<tr>
<th>Excess Return of Portfolio</th>
<th>Equally Weighted</th>
<th>Value Weighted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>S.T.D</td>
</tr>
<tr>
<td>Portfolios Formed on Size</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (Small Size, bottom 30%)</td>
<td>.8959</td>
<td>6.386</td>
</tr>
<tr>
<td>2 (Medium, Middle 40%)</td>
<td>.6913</td>
<td>5.681</td>
</tr>
<tr>
<td>3 (Large Size, Top 30%)</td>
<td>.5631</td>
<td>4.824</td>
</tr>
<tr>
<td>Portfolios Formed on B/M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 (Growth, bottom 30%)</td>
<td>.4429</td>
<td>6.618</td>
</tr>
<tr>
<td>5 (Medium, Middle 40%)</td>
<td>.8688</td>
<td>5.414</td>
</tr>
<tr>
<td>6 (Value, Top 30%)</td>
<td>1.183</td>
<td>5.699</td>
</tr>
<tr>
<td>Portfolios Formed on OP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 (Low Profitability, bottom 30%)</td>
<td>.8039</td>
<td>6.663</td>
</tr>
<tr>
<td>8 (Medium, Middle 40%)</td>
<td>.8345</td>
<td>5.079</td>
</tr>
<tr>
<td>9 (High Profitability, Top 30%)</td>
<td>.8446</td>
<td>5.681</td>
</tr>
</tbody>
</table>
Table 3.2: Smooth coefficient factor loadings of Fama – French five factor model for equally weighted portfolios

<table>
<thead>
<tr>
<th>Factor Loadings</th>
<th>Portfolios Formed on Size</th>
<th>Portfolios Formed on B/M</th>
<th>Portfolios Formed on OP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha</td>
<td>Small Size</td>
<td>Medium</td>
<td>Large Size</td>
</tr>
<tr>
<td>10th Percentile</td>
<td>.0037</td>
<td>-.1669</td>
<td>-.1123</td>
</tr>
<tr>
<td>50th Percentile</td>
<td>.1706</td>
<td>.0104</td>
<td>.0097</td>
</tr>
<tr>
<td>MKT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10th Percentile</td>
<td>.8472</td>
<td>1.030</td>
<td>1.034</td>
</tr>
<tr>
<td>50th Percentile</td>
<td>.8941</td>
<td>1.068</td>
<td>1.063</td>
</tr>
<tr>
<td>90th Percentile</td>
<td>.9155</td>
<td>1.079</td>
<td>1.072</td>
</tr>
<tr>
<td>SMB</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10th Percentile</td>
<td>1.064</td>
<td>.5968</td>
<td>.0246</td>
</tr>
<tr>
<td>50th Percentile</td>
<td>1.087</td>
<td>.6074</td>
<td>.0292</td>
</tr>
<tr>
<td>90th Percentile</td>
<td>1.230</td>
<td>.6646</td>
<td>.0405</td>
</tr>
<tr>
<td>HML</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10th Percentile</td>
<td>.0546</td>
<td>.1066</td>
<td>.0635</td>
</tr>
<tr>
<td>50th Percentile</td>
<td>.2221</td>
<td>.1799</td>
<td>.1420</td>
</tr>
<tr>
<td>90th Percentile</td>
<td>.2561</td>
<td>.2208</td>
<td>.1989</td>
</tr>
<tr>
<td>RMW</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10th Percentile</td>
<td>-.3839</td>
<td>-.2451</td>
<td>-.2699</td>
</tr>
<tr>
<td>50th Percentile</td>
<td>-.2733</td>
<td>-.0380</td>
<td>-.0417</td>
</tr>
<tr>
<td>90th Percentile</td>
<td>-.2181</td>
<td>.1288</td>
<td>.0625</td>
</tr>
<tr>
<td>CMA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10th Percentile</td>
<td>-.3163</td>
<td>-.2767</td>
<td>-.2021</td>
</tr>
<tr>
<td>50th Percentile</td>
<td>-.1768</td>
<td>-.2357</td>
<td>-.1083</td>
</tr>
<tr>
<td>90th Percentile</td>
<td>.0628</td>
<td>-.1511</td>
<td>-.0979</td>
</tr>
</tbody>
</table>
Table 3.3: Smooth coefficient factor loadings of Fama – French five factor model for value weighted portfolios

<table>
<thead>
<tr>
<th>Factor Loadings</th>
<th>Portfolios Formed on Size</th>
<th>Portfolios Formed on B/M</th>
<th>Portfolios Formed on OP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Small Size</td>
<td>Medium</td>
<td>Large Size</td>
</tr>
<tr>
<td>Alpha</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10th Percentile</td>
<td>-.0191</td>
<td>.0019</td>
<td>.0004</td>
</tr>
<tr>
<td>50th Percentile</td>
<td>-.0141</td>
<td>.0241</td>
<td>.0110</td>
</tr>
<tr>
<td>90th Percentile</td>
<td>-.0103</td>
<td>.0518</td>
<td>.0120</td>
</tr>
<tr>
<td>MKT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10th Percentile</td>
<td>.9561</td>
<td>1.031</td>
<td>.9900</td>
</tr>
<tr>
<td>50th Percentile</td>
<td>.9718</td>
<td>1.034</td>
<td>.9930</td>
</tr>
<tr>
<td>90th Percentile</td>
<td>.9747</td>
<td>1.039</td>
<td>.9944</td>
</tr>
<tr>
<td>SMB</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10th Percentile</td>
<td>1.031</td>
<td>.5267</td>
<td>-.1886</td>
</tr>
<tr>
<td>50th Percentile</td>
<td>1.059</td>
<td>.5488</td>
<td>-.1643</td>
</tr>
<tr>
<td>90th Percentile</td>
<td>1.063</td>
<td>.5563</td>
<td>-.1594</td>
</tr>
<tr>
<td>HML</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10th Percentile</td>
<td>.0200</td>
<td>.0302</td>
<td>-.0047</td>
</tr>
<tr>
<td>50th Percentile</td>
<td>.0695</td>
<td>.0692</td>
<td>.0023</td>
</tr>
<tr>
<td>90th Percentile</td>
<td>.0759</td>
<td>.0957</td>
<td>.0137</td>
</tr>
<tr>
<td>RMW</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10th Percentile</td>
<td>-.1739</td>
<td>-.0486</td>
<td>.0154</td>
</tr>
<tr>
<td>50th Percentile</td>
<td>-.1327</td>
<td>.0176</td>
<td>.0195</td>
</tr>
<tr>
<td>90th Percentile</td>
<td>-.1269</td>
<td>.0279</td>
<td>.0371</td>
</tr>
<tr>
<td>CMA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10th Percentile</td>
<td>-.0584</td>
<td>-.0564</td>
<td>-.0053</td>
</tr>
<tr>
<td>50th Percentile</td>
<td>.0098</td>
<td>-.0088</td>
<td>.0023</td>
</tr>
<tr>
<td>90th Percentile</td>
<td>.0509</td>
<td>.0091</td>
<td>.0054</td>
</tr>
</tbody>
</table>
Table 3.4: OLS estimate of Fama – French five factor model for equally weighted portfolios

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>Alpha</th>
<th>MKT</th>
<th>SMB</th>
<th>HML</th>
<th>RMW</th>
<th>CMA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Portfolios Formed on Size</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (Small Size, bottom 30%)</td>
<td>0.1586</td>
<td>0.9055</td>
<td>1.085</td>
<td>0.2398</td>
<td>-0.2358</td>
<td>-0.1670</td>
</tr>
<tr>
<td></td>
<td>(0.0896)</td>
<td>(0.0219)</td>
<td>(0.0307)</td>
<td>(0.0428)</td>
<td>(0.051)</td>
<td>(0.0641)</td>
</tr>
<tr>
<td>2 (Medium, Middle 40%)</td>
<td>0.0148</td>
<td>1.068</td>
<td>0.5954</td>
<td>0.1971</td>
<td>-0.0362</td>
<td>-0.2184</td>
</tr>
<tr>
<td></td>
<td>(0.0476)</td>
<td>(0.0116)</td>
<td>(0.0163)</td>
<td>(0.0228)</td>
<td>(0.0240)</td>
<td>(0.0341)</td>
</tr>
<tr>
<td>3 (Large Size, Top 30%)</td>
<td>0.0162</td>
<td>1.067</td>
<td>0.0131</td>
<td>0.1566</td>
<td>-0.0449</td>
<td>-0.1657</td>
</tr>
<tr>
<td></td>
<td>(0.0408)</td>
<td>(0.0100)</td>
<td>(0.0140)</td>
<td>(0.0195)</td>
<td>(0.0206)</td>
<td>(0.0292)</td>
</tr>
<tr>
<td><strong>Portfolios Formed on B/M</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 (Growth, bottom 30%)</td>
<td>-0.0737</td>
<td>1.034</td>
<td>0.8057</td>
<td>-0.1842</td>
<td>-0.2843</td>
<td>-0.2325</td>
</tr>
<tr>
<td></td>
<td>(0.0752)</td>
<td>(0.0184)</td>
<td>(0.0258)</td>
<td>(0.0350)</td>
<td>(0.0379)</td>
<td>(0.0538)</td>
</tr>
<tr>
<td>5 (Medium, Middle 40%)</td>
<td>0.1017</td>
<td>0.9632</td>
<td>0.7730</td>
<td>0.2883</td>
<td>-0.0017</td>
<td>-0.0313</td>
</tr>
<tr>
<td></td>
<td>(0.0506)</td>
<td>(0.0124)</td>
<td>(0.0174)</td>
<td>(0.0242)</td>
<td>(0.0255)</td>
<td>(0.0362)</td>
</tr>
<tr>
<td>6 (Value, Top 30%)</td>
<td>0.3541</td>
<td>0.9155</td>
<td>0.9152</td>
<td>0.5268</td>
<td>-0.1184</td>
<td>-0.0417</td>
</tr>
<tr>
<td></td>
<td>(0.0730)</td>
<td>(0.0179)</td>
<td>(0.0251)</td>
<td>(0.0349)</td>
<td>(0.0368)</td>
<td>(0.0523)</td>
</tr>
<tr>
<td><strong>Portfolios Formed on OP</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 (Low Profitability, bottom 30%)</td>
<td>0.2113</td>
<td>0.9439</td>
<td>0.9866</td>
<td>0.1531</td>
<td>-0.5423</td>
<td>-0.1558</td>
</tr>
<tr>
<td></td>
<td>(0.0942)</td>
<td>(0.0230)</td>
<td>(0.0323)</td>
<td>(0.0451)</td>
<td>(0.0475)</td>
<td>(0.0674)</td>
</tr>
<tr>
<td>8 (Medium, Middle 40%)</td>
<td>0.0315</td>
<td>0.9582</td>
<td>0.6997</td>
<td>0.2948</td>
<td>0.1983</td>
<td>-0.0150</td>
</tr>
<tr>
<td></td>
<td>(0.0403)</td>
<td>(0.0099)</td>
<td>(0.0139)</td>
<td>(0.0193)</td>
<td>(0.0203)</td>
<td>(0.0289)</td>
</tr>
<tr>
<td>9 (High Profitability, Top 30%)</td>
<td>-0.0322</td>
<td>1.055</td>
<td>0.7957</td>
<td>0.2561</td>
<td>0.3745</td>
<td>-0.1101</td>
</tr>
<tr>
<td></td>
<td>(0.0488)</td>
<td>(0.0119)</td>
<td>(0.0168)</td>
<td>(0.0233)</td>
<td>(0.0246)</td>
<td>(0.0349)</td>
</tr>
</tbody>
</table>
### Table 3.5: OLS estimate of Fama – French five factor model for valued weighted portfolios

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>Alpha</th>
<th>MKT</th>
<th>SMB</th>
<th>HML</th>
<th>RMW</th>
<th>CMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Portfolios Formed on Size</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (Small Size, bottom 30%)</td>
<td>-0.0210</td>
<td>0.9729</td>
<td>1.055</td>
<td>0.0679</td>
<td>-0.1232</td>
<td>0.0095</td>
</tr>
<tr>
<td></td>
<td>(0.0296)</td>
<td>(0.0072)</td>
<td>(0.0102)</td>
<td>(0.0142)</td>
<td>(0.0149)</td>
<td>(0.0212)</td>
</tr>
<tr>
<td>2 (Medium, Middle 40%)</td>
<td>0.0155</td>
<td>1.035</td>
<td>0.5515</td>
<td>0.0687</td>
<td>0.0108</td>
<td>-0.0084</td>
</tr>
<tr>
<td></td>
<td>(0.0326)</td>
<td>(0.0080)</td>
<td>(0.0112)</td>
<td>(0.0156)</td>
<td>(0.0164)</td>
<td>(0.0233)</td>
</tr>
<tr>
<td>3 (Large Size, Top 30%)</td>
<td>0.0142</td>
<td>0.9939</td>
<td>-0.1679</td>
<td>0.0066</td>
<td>0.0215</td>
<td>-0.0004</td>
</tr>
<tr>
<td></td>
<td>(0.0076)</td>
<td>(0.0019)</td>
<td>(0.0026)</td>
<td>(0.0030)</td>
<td>(0.0038)</td>
<td>(0.0054)</td>
</tr>
<tr>
<td>Portfolios Formed on B/M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 (Growth, bottom 30%)</td>
<td>0.0343</td>
<td>1.002</td>
<td>-0.0406</td>
<td>-0.2858</td>
<td>0.1751</td>
<td>-0.0024</td>
</tr>
<tr>
<td></td>
<td>(0.0240)</td>
<td>(0.0059)</td>
<td>(0.0082)</td>
<td>(0.0115)</td>
<td>(0.0121)</td>
<td>(0.0172)</td>
</tr>
<tr>
<td>5 (Medium, Middle 40%)</td>
<td>-0.1143</td>
<td>1.014</td>
<td>-0.0075</td>
<td>0.2688</td>
<td>0.0959</td>
<td>0.1772</td>
</tr>
<tr>
<td></td>
<td>(0.0435)</td>
<td>(0.0106)</td>
<td>(0.0149)</td>
<td>(0.0208)</td>
<td>(0.0219)</td>
<td>(0.0311)</td>
</tr>
<tr>
<td>6 (Value, Top 30%)</td>
<td>-0.0314</td>
<td>1.027</td>
<td>0.1930</td>
<td>0.7619</td>
<td>-0.0387</td>
<td>-0.1037</td>
</tr>
<tr>
<td></td>
<td>(0.0315)</td>
<td>(0.0077)</td>
<td>(0.0108)</td>
<td>(0.0151)</td>
<td>(0.0159)</td>
<td>(0.0225)</td>
</tr>
<tr>
<td>Portfolios Formed on OP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 (Low Profitability, bottom 30%)</td>
<td>0.0352</td>
<td>1.048</td>
<td>0.0727</td>
<td>0.1015</td>
<td>-0.6682</td>
<td>-0.1266</td>
</tr>
<tr>
<td></td>
<td>(0.0283)</td>
<td>(0.0069)</td>
<td>(0.0097)</td>
<td>(0.0136)</td>
<td>(0.0143)</td>
<td>(0.0203)</td>
</tr>
<tr>
<td>8 (Medium, Middle 40%)</td>
<td>-0.0578</td>
<td>0.9957</td>
<td>-0.0584</td>
<td>0.0734</td>
<td>-0.0187</td>
<td>0.0933</td>
</tr>
<tr>
<td></td>
<td>(0.0306)</td>
<td>(0.0075)</td>
<td>(0.0105)</td>
<td>(0.0146)</td>
<td>(0.0154)</td>
<td>(0.0219)</td>
</tr>
<tr>
<td>9 (High Profitability, Top 30%)</td>
<td>0.0467</td>
<td>0.9922</td>
<td>-0.0313</td>
<td>-0.0636</td>
<td>0.3520</td>
<td>-0.0127</td>
</tr>
<tr>
<td></td>
<td>(0.0241)</td>
<td>(0.0059)</td>
<td>(0.0083)</td>
<td>(0.0115)</td>
<td>(0.0121)</td>
<td>(0.0172)</td>
</tr>
</tbody>
</table>
Figure 3.1: Conditional Factor Loadings for equally-weighed size portfolio with respect to real interest rate
Bibliography


