NONPARAMETRIC, STOCHASTIC FRONTIER MODELS WITH MULTIPLE INPUTS AND OUTPUTS

Léopold Simar, Paul Wilson

LIDAM Discussion Paper ISBA
2021 / 03

Nonparametric, Stochastic Frontier Models with Multiple Inputs and Outputs

Léopold Simar   Paul W. Wilson*

August 2021

Abstract

Stochastic frontier models along the lines of Aigner et al. (1977) are widely used to benchmark firms' performances in terms of efficiency. The models are typically fully-parametric, with functional form specifications for the frontier as well as both the noise and the inefficiency processes. Studies such as Kumbhakar et al. (2007) have attempted to relax some of the restrictions in parametric models, but so far all such approaches are limited to a univariate response variable. Some (e.g., Simar and Zelenyuk, 2011; Kuosmanen and Johnson, 2017) have proposed nonparametric estimation of directional distance functions to handle multiple inputs and outputs, raising issues of endogeneity that are either ignored or addressed by imposing restrictive and implausible assumptions. This paper extends nonparametric methods developed by Simar et al. (2017) and Hafner et al. (2018) to allow multiple inputs and outputs in an almost fully-nonparametric framework while avoiding endogeneity problems. We discuss identification issues and properties of the resulting estimators, and examine their finite-sample performance through Monte-Carlo experiments. Practical implementation of the method is illustrated using data on U.S. commercial banks.

Keywords: stochastic frontier, nonparametric, efficiency
JEL classification: C01, C21, C40, C51

---

*Simar: Institut de Statistique, Biostatistique, et Sciences Actuarielles, Université Catholique de Louvain-la-Neuve, Louvain-la-Neuve, Belgium; email: leopold.simar@uclouvain.be.

Wilson: Department of Economics and School of Computing, Clemson University, Clemson, South Carolina, USA; email: pww@clemson.edu.

We are grateful for (i) the Palmetto cluster operated by the Cyber Infrastructure Technology Integration group at Clemson University, and (ii) high-performance computing equipment purchased with funds from U.S. National Science Foundation MRI award no. 1725573. Any remaining errors are solely our responsibility.

1 Introduction and Background

Stochastic frontier models are widely used to analyze the productivity and efficiency of firms, and are attractive because they allow for noise in the data generating process (DGP). The models are typically fully-parametric, with functional form specifications for the frontier as well as both the noise and the inefficiency processes along the lines of the composite-error models of Aigner et al. (1977), Battese and Corra (1977) and Meeusen and van den Broeck (1977). The generic model in this framework is

$$Y = \varphi(X) + \epsilon - \eta, \tag{1.1}$$

where Y is an output quantity, $X \in \mathbb{R}^p_+$ is a vector of input quantities, $\varphi(X)$ is the production frontier, $\eta \geq 0$ is random inefficiency and $\epsilon \in \mathbb{R}$ is random noise with $\mathbb{E}(\epsilon) = 0$.[1] Conditionally on X, η and ϵ are independent and in some studies, both η and ϵ are allowed to be heteroskedastic. Note also that X could be replaced by (X, E), with E denoting environmental factors that may influence the production process. In parametric applications, $\varphi(X)$ is almost always specified as either Cobb-Douglas or Translog. Bauer (1990), Bravo-Ureta and Pinheiro (1993), Coelli (1995) and Kumbhakar et al. (2020) observe that most applied papers specify a normal distribution for the two-sided noise term and a half-normal distribution for the one-sided inefficiency term, i.e., $\epsilon \sim N(0, \sigma^2_\epsilon)$ and $\eta \sim N^+(0, \sigma^2_\eta)$.

This paper presents a highly flexible approach for estimating production efficiency in the presence of multiple inputs and multiple outputs. We propose a flexible, almost fully-nonparametric model based on the directional distance measures of firms' efficiencies developed by Chambers et al. (1996, 1998), but we allow for noise. We then transform the model to facilitate estimation using the ideas of Simar et al. (2017) (hereafter, SVKZ) and Hafner et al. (2018) and to avoid endogeneity issues.[2] The frontier is allowed to be fully

[1] Alternatively, Y might be cost, and X a vector of output quantities and input prices; in this case, the minus sign in front of η becomes a plus sign. Or, Y might be revenue, and X a vector of input quantities and output prices.

[2] As shown in the separate Appendix D.1, our method can also accommodate radial, as opposed to

nonparametric, and we impose only symmetry on the density of the noise term. We specify a specific functional form for the density of the inefficiency term, but we estimate its parameters locally. We show that estimates of marginal rates of substitution and transformation, as well as other features of the original, untransformed model, can be easily recovered from estimates of the transformed model. An empirical illustration is provided by estimating the efficiency of U.S. commercial banks before, during and after the late-2007 financial crisis.

Ours is not the first attempt to move away from the fully-parametric version of (1.1) and the associated restrictive and unrealistic assumptions that are often rejected when tested; see Parmeter and Zelenyuk (2019), Sickles and Zelenyuk (2019) or Kumbhakar et al. (2020) for discussion. A fully nonparametric approach raises identification problems as discussed by Hall and Simar (2002), in addition to possible problems for inference.[3] However, we go farther toward a nonparametric model than other proposed methods, while avoiding a number of problems in other proposals. Fan et al. (1996) and Kuosmanen and Kortelainen (2012) allow the frontier function $\varphi(X)$ to be nonparametric, but the stochastic parts of the their models are still parametric and homoskedastic. The model of Hall and Simar (2002) is fully nonparametric with symmetric noise, but asymptotic identification of the model requires the unrealistic, restrictive assumption that the variance of the noise ϵ goes to zero as the sample size n tends to infinity. Kumbhakar et al. (2007), Simar and Zelenyuk (2011) and Park et al. (2015) use nonparametric local maximum likelihood approaches. In these approaches, both the model for $\varphi(X)$ and the structure of the parameters of the local densities of η and ϵ are functional (nonparametric) and approximated by local polynomials, allowing unspecified heteroskedasticity for both stochastic components. Although this is attractive and flexible, estimation is difficult and computationally expensive.

Recently, Simar et al. (2017) (SVKZ hereafter) suggest a flexible model based also on

directional, distances via a simple modification of the transformation we use below.

[3]Bahadur and Savage (1956) show that even in a very simple setting, inference is not possible without some amount of model structure, i.e., assumptions.

local polynomials, but requiring only local least-squares methods for estimation. This model also allows for heteroskedasticity in both η and ϵ in a functional way, but is much faster to compute and is more robust since estimation is based on moment estimators and the full likelihood is not needed. In particular, only symmetry is required for the density of ϵ. We will adapt and use this approach below.

Other approaches are based on deconvolution techniques. Among these, Daouia et al. (2020) use a Tikhonov regularization technique to solve the ill-posed deconvolution problem, allowing $\varphi(X)$ and η to be fully nonparametric, but requiring the distribution of the noise to be fully known (e.g., normal with zero mean and known variance). Kneip et al. (2015) need only to specify the family (e.g., normal) of the symmetric density for the noise term ϵ, and prove that the model is identified and provide estimates of all the components of the model. A pseudo-likelihood approach is used for estimation, with the density of η approximated by an histogram. Another extension is provided by Florens et al. (2020), where the noise has an unspecified symmetric distribution and the frontier is nonparametric. Identification is obtained by specifying a flexible density for η based on Laguerre polynomials. Both the stochastic noise and inefficiency are allowed to be heteroskedastic and to depend on some extra environmental variables E, but computation is again difficult due to the required highly nonlinear optimization, and results are in many cases numerically unstable.

All of the above approaches are developed for a univariate response Y as in (1.1). Extensions to multivariate Y have also been proposed. Some are based on using polar coordinates in the space of the outputs Y, which is natural for radial efficiency measures. This is the approach of Simar (2007), who extends the univariate approach of Hall and Simar (2002), but the noise has to be of "small size", and Simar and Zelenyuk (2011) who extend the local likelihood approach of Kumbhakar et al. (2007). The latter also provides a stochastic version of traditional envelopment estimators such as data envelopment analysis (DEA) and free-disposal hull (FDH) estimators. In these approaches, the random deviations from the frontier are radial. However, the statistical properties of these approaches have not been established, primarily because they suffer from endogeneity problems that are ignored in each of the papers mentioned above. In each case, after transformation to polar coordinates, the regressors are functions of the endogenous inputs (for the input orientation) or of the endogenous outputs (in the output orientation). As will be seen below, the transformation in our model avoids such issues, and we are able to establish desirable properties (e.g., consistency) of our estimators.

Other approaches are based on modeling directional distance functions. In parametric settings, this involves regressing one of the outputs on the inputs and on the other outputs divided by the dependent variable. This necessarily involves endogeneity issues that are either ignored or solved by adopting restrictive assumptions. For example, Cuesta and Orea (2002) ignore any possible endogeneity while citing Coelli and Perelman (1996), who argue that since ratios of outputs or inputs appear on the right-hand side of the estimated model, exogeneity can plausibly be assumed. Serra et al. (2011) similarly ignore any possible endogeneity. Malikov et al. (2016) estimate a parametric, quadratic, directional distance function which is homogeneous with respect to directions but not with respect to inputs nor outputs, and use a system of equations to deal with potential endogeneity.

Recently Kuosmanen and Johnson (2017) suggest a way to avoid endogeneity when estimating directional distance functions in a nonparametric approach, but their method requires restrictive assumptions on the DGP (as honestly admitted by the authors), including constant returns to scale, non-stochastic frontier points and homoskedasticity of the stochastic components. In addition, the variance of the noise must tend to zero asymptotically as $n \to \infty$ as in Hall and Simar (2002). Moreover, the properties of their proposed estimator, based on convex nonparametric least squares in a first stage, remain unknown.

In contrast to the various proposals for less-than-fully-parametric versions of the model in (1.1), our approach allows for both multiple inputs as well as multiple outputs, is computationally tractable, avoids endogeneity problems, and requires only symmetry for the density of the noise term ϵ. The paper is organized as follows. The next section establishes notation and the basic statistical model as well as its transformation. In Section 3 we discuss identification issues and estimation methods. Section 4 presents empirical results from estimation of the efficiency of U.S. commercial banks, illustrating the practical usefulness of our new methods. Summary and Conclusions are given in Section 5.

Some additional results are given in separate appendices. Links between the two models, as well as how estimates of features of the basic model can be recovered from estimates of the transformed model, are discussed in Appendix A. Appendices B–C briefly discusses other methods that could be used for estimation of our model. Various extensions of our model are discussed in Appendix D, including estimation of radial (i.e., input, output or hyperbolic) distances, inclusion of environmental variables, and how assumptions of free-disposability or convexity can be imposed, and how the choice of direction can be made to be data-driven for purposes of direcitonal distance functions. Section E describes our Monte Carlo experiments and results, providing evidence on the performance of our estimators and methods. Additional details on our empirical application in Section 4 are given in Appendix F.

2 The Statistical Model

2.1 Notation

Let $X \in \mathbb{R}_+^p$ and $Y \in \mathbb{R}_+^q$ denote (random) vectors of input and output quantities, respectively, and let $f(x, y)$ denote the joint density of (X, Y) with support on a strict subset of $\mathbb{R}_+^{p+q}$. Similarly, let $x \in \mathbb{R}_+^p$ and $y \in \mathbb{R}_+^q$ denote fixed, non-stochastic vectors of input and output quantities. In the absence of any noise or uncertainty, the production set

$$\Psi := \{(x, y) \mid x \text{ can produce } y\} \tag{2.1}$$

gives the set of feasible combinations of inputs and outputs, and the support of $f(x, y)$ is a subset of Ψ. In this case, $\Pr((X, Y) \in \Psi) = 1$. But in the presence of noise, the support of $f(x, y)$ may extend outside of Ψ and hence $\Pr((X, Y) \in \Psi) < 1$. Standard assumptions on Ψ in micro-economic theory of the firm include the following.

Assumption 2.1. Ψ *is closed.*

Assumption 2.1 ensures the existence of the efficient frontier, or technology given by

$$\Psi^{\partial} := \{(x, y) \mid (\gamma^{-1}x, \gamma y) \notin \Psi \text{ for } \gamma > 1\}. \tag{2.2}$$

In cases where $q = 1$, the loci of points in Ψ^{∂} comprise a function $\varphi(x)$, i.e., the frontier production function.

Let $p \geq 1$, $q \geq 1$ and define $r := p + q$. Define a fixed, non-stochastic, $(r \times 1)$ direction vector $d = (-d_x', d_y')'$ where $d_x \in \mathbb{R}_+^p$ and $d_y \in \mathbb{R}_+^q$.[4] Then for a fixed point $(x, y) \in \mathbb{R}_+^r$, distance to the frontier in the direction d is measured by the directional distance function

$$\delta(x, y \mid d) := \sup\{\delta \mid (x - \delta d_x, y + \delta d_y) \in \Psi\} \tag{2.3}$$

proposed by Chambers et al. (1996, 1998) whenever $\delta(x, y \mid d)$ exists.[5]

In the absence of noise, properties of nonparametric DEA and FDH estimators of the distance function defined in (2.3) are well-developed; see Kneip et al. (2008) and the sources cited therein or see Simar and Wilson (2013, 2015) for recent surveys. Except for the required assumption of no noise, DEA estimators (which assume strong disposability of

[4] Several papers that use directional distance functions (e.g., Vardanyan and Noh, 2006; Atkinson and Tsionas, 2016; and Färe et al., 2017) note that the chosen direction can influence the magnitudes and rankings of estimated inefficiencies. This is well-known; Wheelock and Wilson (2008) and Wilson (2011) suggest using hyperbolic, radial distances rather than input or output radial distances for this reason. Some who model the directional distance function have allowed the directions to be determined by data. The usual criterion involves choosing directions to minimize estimate inefficiency; example include Atkinson and Tsionas (2016) and Färe et al. (2017). Layer et al. (2020) propose setting directions equal to medians of the data after looking at results from some Monte Carlo simulations. Many papers use means of the input and output variables to determine directions. Atkinson and Tsionas (2016) and Färe et al. (2017) require substantially more parametric structure than we impose here.

[5] If the path $(x, y) + \delta d$ does not pass through Ψ for any real-valued δ, then $\delta(x, y \mid d)$ does not exist.

inputs and outputs and convexity of Ψ) and FDH estimators (which only assume strong disposability of inputs and outputs) are very flexible, require no functional form assumptions, and easily handle multiple outputs and multiple inputs. DEA and FDH estimators rely on the assumption that the support of the random variables (X, Y) coincides with a subset of Ψ.[6]

Regardless of whether noise is present, the production set Ψ, its frontier Ψ^∂, and the distances defined in (2.3) are unobserved and must be estimated from data. Some additional assumptions are needed to define a statistical model. We first introduce a basic model below in Section 2.2, and then transform this model in Section 2.3 to facilitate estimation. As will be seen, the basic model amounts to a system of equation describing the underlying process that generates the observed data, while the transformed model is used for estimation.

2.2 The Basic Model

We first assume that the underlying theoretical economic process generates exogenous, efficient, but latent, production plans.

Assumption 2.2. *The Data Generating Process (DGP) generates random, optimal production plans that belongs to the efficient boundary Ψ^∂ via a well-defined probability mechanism, providing identically, independently distributed (iid) values $W_i^\partial := (X_i^\partial, Y_i^\partial)$, $i = 1, \ldots, n$.*

The optimal plans W_i^∂ (i.e., frontier points) are unobserved due to inefficiency and noise. Now consider a given, fixed direction vector d. The earlier comments in footnote 4 notwithstanding, empirical researchers typically choose a fixed direction vector d (e.g., see the discussion by Färe et al., 2008, p. 534), often with elements equal to means or medians of the corresponding variables. For our purposes, the point is to make d fixed, non-stochastic and the same for all firms (the latter is sometimes called the "egalitarian" evaluation in (2.3)). Let observed input-output pairs be denoted by $W_i := (X_i, Y_i)$.

[6]For additional technical details, see Kneip et al. (2008).

Assumption 2.3. *Given a direction vector d, the set of sample observations $\mathcal{S}_n = \{W_i\}_{i=1}^n$ is generated according the model $W_i = W_i^\partial + \xi_i d$, i.e., component by component*

$$\begin{bmatrix} X_i \\ Y_i \end{bmatrix} = \begin{bmatrix} X_i^\partial \\ Y_i^\partial \end{bmatrix} + \xi_i d, \tag{2.4}$$

where the ξ_i are, conditionally on W_i^∂, independent, scalar random variables whose characteristics may depend on W_i^∂.

Both the inefficiency and the noise processes (2.4) operate on both inputs and outputs, as is standard in the literature on directional distance functions. Assumptions 2.1–2.4 ensure we have a random sample $\mathcal{S}_n = \{(X_i, Y_i)\}$ of size n with the latent efficient variables W_i^∂ playing the role of unobserved, exogenous variables. The components ϵ_i and η_i of the independent ξ_i may be heteroskedastic. These assumptions induce the joint density $f(x, y)$ of the observables (X, Y).

Next, we decompose ξ into additive inefficiency and noise terms to complete the assumptions of our basic model.

Assumption 2.4. *The stochastic term ξ_i in Assumption 2.3 is the convolution of an inefficiency component $\eta_i \in \mathbb{R}_+$ and a noise component $\epsilon_i \in \mathbb{R}$, i.e.*

$$\xi_i := \epsilon_i - \eta_i, \tag{2.5}$$

where conditionally on W_i^∂, ϵ_i and η_i are independent, with

$$\epsilon_i \mid W_i^\partial \sim Sym\left(0, \widetilde{\sigma}_\epsilon(W_i^\partial)\right) \tag{2.6}$$

and

$$\eta_i \mid W_i^\partial \sim D_+\left(\widetilde{\sigma}_\eta(W_i^\partial)\right) \tag{2.7}$$

where $Sym(0, a)$ denotes a symmetric (around zero), two-sided distribution on $\mathbb{R}$ with standard deviation $a \geq 0$, and $D_+(b)$ denotes a one-sided distribution on $\mathbb{R}_+$ belonging to some *one-parameter scale family with parameter* $b \geq 0$.[7]

As an aside, consider the following special case of our basic model. Suppose that $q = 1$ (univariate output) and that the chosen direction is output oriented with $d = (0_p', 1)'$ where 0_p is a p-vector of zeros. Then (2.4) simplifies to

$$\begin{aligned} X_i &= X_i^{\partial}, \\ Y_i &= Y_i^{\partial} + \epsilon_i - \eta_i. \end{aligned} \tag{2.8}$$

Since $X_i = X_i^{\partial}$ is observed (because the corresponding elements of d are zero), and Y_i is univariate, we can in this case describe the frontier points by the production function $\mathbb{R}_+^p \mapsto \mathbb{R}_+^1$ such that $Y_i^{\partial} = Y^{\partial}(X_i)$. By conditioning on X_i, the model in (2.8) can be written as

$$Y_i = Y^{\partial}(X_i) + \epsilon_i - \eta_i, \tag{2.9}$$

where $\epsilon_i \mid X_i \sim \text{Sym}(0, \sigma_\epsilon(X_i))$ and $\eta_i \mid X_i \sim \text{D}_+(\sigma_\eta(X_i))$ since the parameters of the scale functions in (2.6)–(2.7) can now be expressed as functions of X_i. The model in (2.9) is identical to the output-oriented model suggested by Simar et al. (2017) for the case of a univariate output Y. Kumbhakar et al. (2007) use the same model with a specific localized distribution (i.e., Gaussian) for ϵ_i in order to derive the local likelihood function. Moreover, it is obvious that (2.9) nests the fully-parametric models of Aigner et al. (1977), Battese and Corra (1977) and Meeusen and van den Broeck (1977) as a special case, but allows far more flexibility than the fully-parametric models. In addition, (2.9) is a special case of our basic model in (2.4), which allows $q > 1$.

2.3 The Transformed Model

In order to estimate the frontier as well as features of the inefficiency process, we translate the problem to a new coordinate system to obtain a transformed version of the basic model

[7] A density $f(\cdot)$ belongs to a one-parameter scale family if it can be written as $f(\cdot) = (1/\sigma)\tilde{f}(\cdot/\sigma)$ for some $\sigma > 0$, where $\tilde{f}(\cdot)$ is any density on $\mathbb{R}_+$. We will call $\tilde{f}$ the *reference* density of the family. Examples include the Half-Normal and Exponential distributions, and the Gamma and Weibull distributions with fixed shape parameters.

defined by Assumptions 2.1–2.4. First, consider an arbitrary, but fixed, orthonormal basis $V_d = \begin{bmatrix} v_1 & \dots & v_{r-1} \end{bmatrix}$ for the direction vector d. Then V_d is an $r \times (r-1)$ matrix, with $v_j'v_j = 1$, $v_j'v_k = 0$ for $j \neq k$, and $v_j'd = 0$. Consequently, $V_d'V_d = I_{r-1}$ where I_{r-1} denotes the $(r-1) \times (r-1)$ identity matrix. Moreover, V_d depends only on the given direction vector d. [8] Using the matrix V_d, define the $(r \times r)$ rotation matrix

$$R_d = \begin{bmatrix} V_d' \\ d'/||d|| \end{bmatrix} \tag{2.10}$$

whose transpose is $R_d' = \begin{bmatrix} V_d & d/||d|| \end{bmatrix}$ and where $||\cdot||$ denotes the L_2-norm. Clearly, R_d is an orthogonal matrix. Hence $R_d'R_d = I_r$, implying $R_d^{-1} = R_d'$. Moreover, by construction, $R_d d = \begin{bmatrix} 0_{r-1}' & ||d|| \end{bmatrix}'$. It is important to remember that the rotation matrix R_d is specific to the given direction vector d as signified by the subscript "d". If the direction vector is changed, this will affect the rotation matrix as well as everything that follows. But as with V_d, R_d depends only upon the given direction vector d.

In the special case of a univariate output ($q = 1$) with output-oriented direction $d = (0_p', 1)'$ considered above in (2.8), it is easy to show that the matrix R_d is given by the identity matrix of order r. In this case, no transformation of the basic model is needed, since the setup is the same as that of SVKZ. Otherwise, however, we can transform the observations $W_i = (X_i, Y_i) \in \mathcal{S}_n$ to a new coordinate system by applying the rotation matrix R_d in (2.10) to obtain

$$\begin{bmatrix} Z_i \\ U_i \end{bmatrix} = R_d W_i = R_d \begin{bmatrix} X_i \\ Y_i \end{bmatrix}, \tag{2.11}$$

where $Z_i \in \mathbb{R}^{r-1}$ and $U_i \in \mathbb{R}^1$ for each $i = 1, \dots, n$. Moreover, the independent, identically distributed (iid) sample $\mathcal{S}_n$ is transformed to an iid sample $\mathcal{S}_n(d) = \{(Z_i, U_i)\}_{i=1}^n$. Obviously, the (Z_i, U_i) depend on the direction vector d through the rotation matrix R_d, but we avoid

[8] Note that the matrix V_d is not unique. Let Υ_{r-1} be any $(r-1)$ orthogonal matrix, then it is easy to see that $\widetilde{V}_d = V_d\Upsilon_{r-1}$ is also an orthonormal basis for d. This does not create any problem for what follows, provided V_d is fixed after it is selected. If two applied researchers using different algorithms to build the matrix obtain matrices $V_d \neq \widetilde{V}_d$, the link between the two matrices is given by the orthogonal matrix $\Upsilon_{r-1} = V_d'\widetilde{V}_d$.

further complicating the notation. Note that $U_i = d'W_i/||d||$ and hence is invariant to the ordering of the inputs and outputs since the ordering of the elements of the direction vector d necessarily corresponds to whatever order of the elements of X_i and Y_i is chosen.

From a geometrical point of view, we have a linear transformation from $w = (x, y) \in \mathbb{R}^r$ to $(z, u) \in \mathbb{R}^r$, where $\begin{bmatrix} z' & u \end{bmatrix}' = R_d w$. This one-to-one transformation can be inverted by writing $w = R_d' \begin{bmatrix} z' & u \end{bmatrix}'$. This is illustrated in the case $p = q = 1$ in Figure 1 where the origin in the new (z, u)-space corresponds to the origin in the original (x, y)-space. The rotation transforms the production set Ψ in (x, y)-space to the set

$$\Gamma_d = \left\{ (z, u) \in \mathbb{R}^r \mid R_d' \begin{bmatrix} z' & u \end{bmatrix}' \in \Psi \right\} \tag{2.12}$$

in (z, u)-space. Moreover, the density $f(x, y)$ is transformed to a new density $g(z, u) = g(u \mid z)g(z)$. Since u is in the direction d, and the $r - 1$ coordinates z are orthogonal to d, the frontier Ψ^∂ in (x, y)-space is transformed to the function $\phi(z)\colon \mathbb{R}_+^{r-1} \mapsto \mathbb{R}^1$ in (z, u)-space such that

$$\phi(z) = \sup \{ u \mid (z, u) \in \Gamma_d \} . \tag{2.13}$$

In (x, y)-space, the frontier is an $(r - 1)$-manifold, while in (z, u) space it becomes a scalar-valued, $(r - 1)$-variate function $\phi(z)$. So we have

$$\Gamma_d = \{ (z, u) \in \mathbb{R}^r \mid u \leq \phi(z) \} . \tag{2.14}$$

The frontier points $(X_i^\partial, Y_i^\partial)$ have coordinates (Z_i, U_i^∂) in (z, u)-space, and the observations (X_i, Y_i) have coordinates (Z_i, U_i), as illustrated in Figure 1. For each $i = 1, \ldots, n$, the random deviations from $(X_i^\partial, Y_i^\partial)$ along the direction d to observations (X_i, Y_i) correspond to random deviations from (Z_i, U_i^∂) to the observations (Z_i, U_i) in (z, u)-space. The heteroskedastic nature of ξ_i is now captured by the values of Z_i.

Pre-multiplying both sides of the basic model in (2.4) by R_d yields the transformed model

$$\begin{bmatrix} Z_i \\ U_i \end{bmatrix} = \begin{bmatrix} Z_i^\partial \\ U_i^\partial \end{bmatrix} + \xi_i \begin{bmatrix} 0_{r-1} \\ ||d|| \end{bmatrix} , \tag{2.15}$$

Then the transformed model in (2.15) can be written equivalently as

$$\begin{aligned} Z_i &= Z_i^\partial, \\ U_i &= U_i^\partial + ||d||\epsilon_i - ||d||\eta_i, \end{aligned} \tag{2.16}$$

where the equation for U_i is univariate. Note that (2.16) has structure analogous to that of the special case of our basic model described by (2.8). Since $Z_i = Z_i^\partial$ is observed, we can rewrite the frontier points in the (z,u)-space as $(Z_i, U^\partial(Z_i)) = (Z_i, \phi(Z_i))$. Therefore, by conditioning on Z_i, the second equation of (2.16) can be written as

$$U_i = \phi(Z_i) + ||d||\epsilon_i - ||d||\eta_i, \tag{2.17}$$

where the scale functions $\widetilde{\sigma}_\epsilon(W_i^\partial)$ and $\widetilde{\sigma}_\eta(W_i^\partial)$ in ξ_i can be expressed as functions of Z_i since W_i^∂ can be expressed as a function of Z_i only. Hence the heteroskedastic features defined in (2.6)–(2.7) of Assumption 2.4 can be written as

$$\epsilon_i \mid Z_i \sim \mathrm{Sym}(0, \sigma_\epsilon(Z_i)) \tag{2.18}$$

and

$$\eta_i \mid Z_i \sim \mathrm{D}_+(\sigma_\eta(Z_i)). \tag{2.19}$$

It is important to note that

$$\begin{bmatrix} Z_i^\partial \\ U_i^\partial \end{bmatrix} = R_d \begin{bmatrix} X_i^\partial \\ Y_i^\partial \end{bmatrix} = R_d W_i^\partial. \tag{2.20}$$

The first line of (2.16) results from the facts that (i) R_d depends only upon the fixed d, and (ii) $R_d d = \begin{bmatrix} 0'_{r-1} & 1 \end{bmatrix}'$ by construction as discussed above. Then by the first line of (2.16), Z_i is exogenous since W_i^∂, and hence Z_i^∂, are exogenous. In the transformed model (2.17), only the scalar left-hand side variable U_i is endogenous. This greatly simplifies estimation, and stands in sharp contrast to most of the literature where researchers work in the original (x,y)-space where all of the elements of X_i and Y_i are potentially endogenous. Of course this depends on a fixed, exogenously-chosen direction vector, but this is in fact what most empirical researchers adopt.[9]

In order to identify the frontier function $\phi(Z_i)$ in (2.13), a functional form for the distribution D_+ in 2.19 must be specified. Various possibilities are mentioned above in footnote 7; for purposes of illustration, we adopt the following.

Assumption 2.5. *The distribution of η, conditional on Z, is half-normal with scale parameter $\sigma_\eta(Z)$, i.e.,*

$$\eta \mid Z \sim N^+ \left(0, \sigma_\eta^+(Z)\right). \tag{2.21}$$

In addition, for purposes of estimation using local-linear methods along the lines of SVKZ, two additional assumptions are needed.

Assumption 2.6. *For any orthonormal basis V_d of d, the function $\phi(z)$ defined in* (2.13) *is differentiable at all points z in its domain.*

Assumption 2.7. *For all $i \in \{1, 2, \dots\}$, the conditional moments $\mathbb{E}\left(|U_i|^{2+\nu} \mid Z_i\right)$ $\mathbb{E}\left(|(\xi_i + \mu_\eta(Z_i))^3|^{2+\nu^*} \mid Z_i\right)$ exist for some $\nu > 0$ and $\nu^* > 0$.*

For a given Z_i, there is a unique U_i^∂ and W_i^∂ in the direction d. By Assumption 2.4, $\mathbb{E}(\eta_i^j \mid Z_i) = k_j \sigma_\eta^j(Z_i)$ for all $j \geq 1$, provided k_j, the jth moment of the reference density of the family $\tilde{f}$ exists (see the remarks in footnote 7). Moreover, Assumption 2.4 ensures that the conditional mean efficiency function is given by

$$\mu_\eta(Z_i) := \mathbb{E}(\eta_i \mid Z_i) = k_1 \sigma_\eta(Z_i). \tag{2.22}$$

Under Assumption 2.5, $k_1 = \sqrt{2/\pi}$. Replacing Assumption 2.5 with a different distributional assumption will change the value of k_1. Assumption 2.6 imposes some smoothness on the frontier, in all directions, while Assumption 2.7 imposes some minor restrictions on the tail

[9] In the separate Appendix D.4, we discuss how our estimation methods described below in Section 3 could be extended to allow for a data-driven choice of the direction vector, although the computational burden would increase significantly over what we propose below.

behavior of the distributions of ϵ_i and η_i in (2.18)–(2.19). Both assumptions are needed to ensure consistency of the local-linear estimators used below in Section 3.

In the next section, we show how the unknown functions $\phi(Z_i)$, $\sigma_\epsilon^2(Z_i)$ and $\sigma_\eta(Z_i)$ can be easily estimated using the transformed model in (2.17) using using standard nonparametric regression estimators (e.g., local polynomial or orthogonal series estimators). In particular the "σ functions" in (2.18)–(2.19) will be functional parameters, enhancing the flexibility of the model. Estimates of the conditional mean efficiency function in (2.22) are obtained directly from the transformed model. Estimates of features of the basic model, such as marginal rates of substitution or transformation, as well as derivatives of conditional mean efficiency with respect to inputs or outputs, can be recovered easily from estimates of the transformed model due to the fact that the transformation from (x, y)-space to (z, u)-space is one-to-one. Details are given in the separate Appendix A.

3 Estimation and Inference

3.1 Estimation Methodology and Basic Properties

As discussed above in Section 2.3, the transformed model given by (2.17)–(2.19) involves a scalar response variable U_i, a link function $\phi(Z_i)$, and a composite error term $\xi_i = \epsilon_i - \eta_i$. The distributions of $\epsilon_i \in \mathbb{R}$ and $\eta_i \in \mathbb{R}_+$, conditional on Z_i, are given in (2.18)–(2.19). These are necessarily related to the direction vector d through the dependence on Z_i. It is easy to prove that the model is identified once the reference family is chosen (as in Assumption 2.5) using the sufficient condition for identification given by Florens et al. (2020).

We use the moments-based method of SVKZ and its extension described by Hafner et al. (2018) (hereafter, HMS) to estimate our transformed model. Alternative methods that might be used are discussed briefly in the separate Appendix B. However, our stochastic specification is quite flexible since it is localized (i.e., the specification may change with the value of z), and the convolution of a one-parameter scale family with a symmetric density is also quite flexible. This is illustrated in Figure 2 for the particular choice of the one-parameter scale density in Assumption 2.5, with various values of the scale parameter σ_η and various symmetric densities for ϵ (Normal, Laplace, Uniform and a Mixture of two Normal densities). Figure 2 shows that the density of the convolution of noise and efficiency can have a wide variety of shapes under our assumptions. Another advantage of our estimation approach is its numerical easiness and relatively low computational requirements, since only local least-squares methods are used and most of the relevant information about the model can be recovered by assuming only symmetry of ϵ and selecting a specific one-parameter scale family for the distribution of efficiency. Asymptotic properties of our estimators are derived by SVKZ and HMS.

We first show how the SVKZ estimation method can be used to estimate the transformed model given by (2.17)–(2.19). The method can be viewed as a nonparametric version of the modified OLS procedure suggested by Richmond (1974) that is sometimes used in fully parametric settings (e.g., see Lovell, 1993 for discussion). The SVKZ approach allows the link function $\phi(z)$ to remain nonparametric and uses only information of the first three local moments of the random terms. Consequently, the method used here is less sensitive to error in specifications of the local densities of the stochastic terms than some of the approaches discussed in the separate Appendix B.

To explain the procedure, note that $\mu_\eta(Z_i)$ defined in (2.22) must be nonnegative for all $i = 1, \ldots, n$. To simplify notation, assume without loss of generality in all that follows that the direction vector d has been normalized so that $||d|| = 1$. Define $\varepsilon_i := \xi_i + \mu_\eta(Z_i)$ and $r_1(Z_i) := \phi(Z_i) - \mu_\eta(Z_i)$. Clearly, $\mathbb{E}(\varepsilon_i \mid Z_i) = 0$. Then necessarily

$$U_i = r_1(Z_i) + \varepsilon_i \tag{3.1}$$

and $\mathbb{E}(U_i \mid Z_i) = r_1(Z_i) = \phi(Z_i) - \mu_\eta(Z_i)$. The conditional mean function $r_1(Z_i)$ can be viewed as an average production function that can be estimated using standard nonparametric regression estimators (e.g., the local-linear estimator).

The assumption of symmetry of $\epsilon_i \mid Z_i$ ensures that

$$\mathbb{E}(\varepsilon \mid z) = 0, \tag{3.2}$$

$$\mathbb{E}(\varepsilon^2 \mid z) = \mathrm{VAR}(\epsilon \mid z) + \mathrm{VAR}(\eta \mid z) > 0 \tag{3.3}$$

and

$$\mathbb{E}(\varepsilon^3 \mid z) = -\mathbb{E}\left[(\eta - \mu_\eta(z))^3 \mid z\right]. \tag{3.4}$$

Now let $r_j(z) := \mathbb{E}(\varepsilon^j \mid z)$ for $j \in \{2,\ 3\}$. Under Assumptions 2.4 and 2.5, it is straightforward to show that

$$\mu_\eta(z) = \mathbb{E}(\eta \mid z) = \sqrt{\frac{2}{\pi}}\sigma_\eta(z), \tag{3.5}$$

$$r_2(z) = \sigma_\epsilon(z)^2 + \frac{\pi - 2}{\pi}\sigma_\eta(z)^2 \tag{3.6}$$

and

$$r_3(z) = \sqrt{\frac{2}{\pi}}\frac{\pi - 4}{\pi}\sigma_\eta^3(z) \leq 0. \tag{3.7}$$

From (3.7) is is clear that the the scale function $\sigma_\eta(z)$ is identified by the third (local) moment $r_3(z)$, and this in turn provides identification of $\mu_\eta(z)$.

Estimation is straightforward. First, (3.1) is estimated by local linear least squares (LLLS), yielding estimates $\widehat{r}_1(z)$ and $\widehat{\nabla r_1}(z)$ of $r_1(z)$ and $\partial r_1(z)/\partial z$, respectively. The properties of these estimators are well-known; e.g., see Fan and Gijbels (1996) or Li and Racine (2007). Second, compute the residuals $\widehat{\varepsilon}_i = U_i - \widehat{r}_1(Z_i)$ for $i = 1,\ \ldots,\ n$, and again using LLLS, regress $\widehat{\epsilon}_i^2$ and $\widehat{\epsilon}_i^3$ on the Z_i, $i = 1,\ \ldots,\ n$ to obtain corresponding estimates $\widehat{r}_j(Z_i)$, $j \in \{1,\ 2\}$. SVKZ show that the usual asymptotic properties (i.e., rates of convergence and asymptotic normality) for $\widehat{r}_j(z)$ for $j = 2, 3$ hold under mild regularity conditions. In particular, if bandwidths for the r components of the Z_i have the same order, then the $\widehat{r}_j(z)$, $j \in \{2,\ 3\}$ converge at rate $(nh^{p+q-1})^{1/2}$ where the bandwidth h is written without a subscript since all are of the same order. It is well-known that the optimal order of h for local-linear regression is $h = O\left(n^{-1/(p+q+3)}\right)$, which leads to an optimal rate $n^{2/(p+q+3)}$ for the $r_j(z)$. Here, as elsewhere with nonparametric regression, there is no escape from the curse of dimensionality, but there is minimal risk of specification error.[10]

Substituting $\widehat{r}_3(z)$ into (3.7) and rearranging terms yields the estimator

$$\widehat{\sigma}_\eta(z) = \max\left\{0,\ \left[\sqrt{\frac{\pi}{2}}\frac{\pi}{\pi-4}\widehat{r}_3(z)\right]^{1/3}\right\} \geq 0 \tag{3.8}$$

of the scale function $\sigma_\eta(z)$. The max operator is used to ensure that the estimate is non-negative; this is discussed further in Section 3.2. Substituting $\widehat{\sigma}_\eta(z)$ this into (3.5) yields an estimator of mean conditional inefficiency $\mu_\eta(z)$, i.e.,

$$\widehat{\mu}_\eta(z) = \sqrt{\frac{2}{\pi}}\widehat{\sigma}_\eta(z). \tag{3.9}$$

Similarly, an estimator

$$\widehat{\sigma}^2_\epsilon(z) = \max\left\{0,\ \widehat{r}_2(z) - \frac{\pi-2}{\pi}\widehat{\sigma}^2_\eta(z)\right\} \geq 0 \tag{3.10}$$

of the variance of ϵ is obtained from (3.6), although this is not needed if one is only interested in estimates of the frontier and of mean conditional inefficiency. Finally, an estimator of the frontier function in the transformed space is given by

$$\widehat{\phi}(z) = \widehat{r}_1(z) + \widehat{\mu}_\eta(z). \tag{3.11}$$

In addition, the partial derivative estimates $\widehat{\nabla r_j}(z)$, $j = 1, \ldots, 3$ can be used to recover estimates $\widehat{\nabla\mu_\eta}(z)$ and $\widehat{\nabla\phi}(z)$ of the partial derivatives of $\mu_\eta(z)$ and $\phi(z)$ with respect to the r elements of z. These estimates have the usual asymptotic properties for derivatives estimated by LLLS (see, for discussion, Li and Racine, 2007, Section 2.4 and SVKZ, Section

[10]SVKZ under-smooth by choosing a slightly smaller order for h to avoid the asymptotic bias in the limiting normal distribution. SVKZ explain in their Appendix that the bandwidths should satisfy $nh^{p+1-1}(\log n)^{-4} \to \infty$ and $nh^{p+q+3} \to 0$ as $n \to \infty$. Hence the order of the bandwidths should be $O(n^{-(1+\upsilon)/(p+q+3)})$ for some $\upsilon > 0$.

4.1, where estimators of elasticities of $\mu_\eta(z)$ with respect to z and their asymptotic properties are derived).

One can also derive predictions of individual efficiencies for each observation, using a local version of the Jondrow et al. (1982) approach. However, this requires specification of a particular density in the family $\text{Sym}(0, \sigma_\epsilon(z))$ (e.g., normality). Then the frontier estimates $\widehat{\phi}(Z_i)$ can be used to compute residuals $\xi_i = U_i - \widehat{\phi}(Z_i)$, which are in turn used to compute $\widehat{\eta}_i = \mathbb{E}(\eta_i \mid \widehat{\xi}_i, Z_i)$ along the lines of Jondrow et al. (1982). However, this should be used with caution. As observed by SVKZ, the resulting $\widehat{\eta}_i$ are predicted values based on *a single, particular* realization $\widehat{\xi}_i$. In fact, the $\widehat{\eta}_i$ are not predictions, but rather plug-in estimates of the conditional mean $\mathbb{E}(\eta_i \mid \xi_i, Z_i)$ as discussed by Simar and Wilson (2010) and Wheat et al. (2014). Even so, confidence intervals for $\widehat{\eta}_i$ are likely to be quite wide since the estimate is based on a single observation. Moreover, if one is interested in predictions, one should estimate prediction intervals as discussed by Simar and Wilson (2010) instead of confidence intervals for the conditional mean.

3.2 Handling the "Wrong Skewness" Problem

It is well-known that that the convolution of an unbounded, symmetrically-distributed random variable and a one-sided random variable may yield finite samples that are skewed in either direction. In both the basic and transformed models discussed above, one might expect the $\{\xi_i\}_{i=1}^n$ to be skewed to the left, but skewness to the right occurs with some probability depending on the noise-to-signal ratio $\sqrt{\text{VAR}(\epsilon)/\text{VAR}(\eta)}$. In samples of sizes often encountered in empirical work, and with reasonable noise-to-signal ratios, this phenomenon may occur with high probability (e.g., see Simar and Wilson, 2010, Table 1). This is sometimes called the "wrong skewness" problem, but as discussed by Simar and Wilson (2010), it can occur in the absence of any specification error or other mistakes; i.e., the problem may arise in cases there is nothing "wrong" with the model nor with anything else. In the context of our model, $r_3(z) \leq 0$ by construction, but in finite samples, one may find $\widehat{r}_3(z) > 0$. Since local estimation is used here, the effective sample size is of order $n^{4/(p+q+3)}$ due to the use of bandwidths with optimal order. Hence the problem is even more serious here than in the fully-parametric context discussed by Simar and Wilson (2010).

The SVKZ estimator deals with unexpected skewness by simply truncating the estimator $\widehat{\sigma}_\eta(z)$ at zero in (3.8). Alternatively, one might consider imposing the constraint $r_3(z) \leq 0$ in the local linear regression of the $\widehat{\xi}_i^3$ on Z_i as mentioned by SVKZ. Or, instead one might use the local exponential estimator proposed by Ziegelmann (2002). As discussed in the separate Appendix C, constrained local linear estimation does not change the truncation in (3.8), but affects estimation of derivatives. As also discussed in Appendix C, local exponential estimation may be attractive from a theoretical viewpoint, but it loses the numerical flexibility of local linear methods and its computational burden is considerably larger than local linear methods. A third approach, adopted here, is to extend the ideas developed by HMS for fully parametric settings to the localized framework here.[11]

The idea is straightforward, and we provide only a brief sketch of the idea here. Details on derivations are given in HMS. To illustrate, we maintain the local half-normal specification for the distribution of the inefficiency. Let k_j denote the jth raw moment of the reference density $N^+(0,1)$ (recall footnote 7). We have

$$E\left[(\eta - \mu_\eta(z))^3 \mid z\right] = a_3^+ \gamma^3(z) > 0 \tag{3.12}$$

where $a_3^+ = k_3 - 3k_1k_2 + 2k_1^3$. In the "regular" case, $\gamma(z) = \sigma_\eta(z) > 0$ and the skewness of $\eta \mid z$ is positive. Denote the density of $\eta \mid z$ in this case by $h^+(\eta \mid z) = \frac{1}{\gamma(z)}\widetilde{h}^+\left(\frac{\eta}{\gamma(z)}\right)$.

Now consider the density $h^-(\eta|z)$ with support on $\mathbb{R}^+$ having the same mean as h^+, but with negative skewness. This is obtained by reflecting the density h^+ around zero, shifting

[11] Of course, the "wrong skewness" problems goes away asymptotically, but empirical researchers are typically confronted with samples of fixed sizes. All three methods mentioned here are asymptotically equivalent, since the unexpected skewness disappears with probability one as $n \to \infty$, but they may differ in finite samples.

so that the support lies in $\mathbb{R}^+$, and then truncating on the right so that the resulting density has the same mean as h^+. This yields

$$h^-(\eta \mid z) = \frac{1}{|\gamma(z)|}\widetilde{h}^-\left(\frac{\eta}{|\gamma(z)|}\right), \tag{3.13}$$

where $\widetilde{h}^- \neq \widetilde{h}^+$ (see HMS for an explicit expression for $\widetilde{h}^-$ and for its properties; we use use only its first 3 moments). By construction, the mean given by the density $h^-(\eta|z)$ is equal to $\mu_\eta(z) \geq 0$ (same mean as $h^+(\eta|z)$), for $\eta \mid z$ drawn from $h^-(z)$,

$$E\left[(\eta - \mu_\eta(z))^3 \mid z\right] = a_3^-\gamma(z)^3 < 0, \tag{3.14}$$

where $\gamma(z) < 0$ and $\sigma_\eta(z) = |\gamma(z)| > 0$ For the half-normal family, $a_3^+ = \sqrt{2/\pi}\,((4-\pi)/\pi)$ and $a_3^- \approx 0.016741474$ (again, see HMS for further details).

Adapting equations (37)–(38) in HMS to our local version, we have

$$\widehat{\sigma}_\eta(z) = \begin{cases} \left[-\frac{1}{a_3^+}\widehat{r}_3(z)\right]^{1/3} & \text{if } \widehat{r}_3(z) \leq 0; \\ \left[\frac{1}{a_3^-}\widehat{r}_3(z)\right]^{1/3} & \text{otherwise.} \end{cases} \tag{3.15}$$

Asymptotically, one should expect $\widehat{r}_3(z) \leq 0$, so the asymptotic properties are the same as for the SVKZ estimator. But in finite samples, the HMS estimator in (3.15) avoids excessive numbers of estimates equal to zero.

Similarly, if an estimator of the variance of ϵ is required, using the half-normal family for the inefficiency term we have

$$\widehat{\sigma}^2_\epsilon(z) = \begin{cases} \max\left[0,\ \widehat{r}_2(z) - a_2^+\widehat{\sigma}^2_\eta(z)\right] & \text{if } \widehat{r}_3(z) \leq 0; \\ \max\left[0,\ \widehat{r}_2(z) - a_2^-\widehat{\sigma}^2_\eta(z)\right] & \text{otherwise,} \end{cases} \tag{3.16}$$

where $a_2^+ = (\pi - 2)/\pi$ and $a_2^- \approx 0.14471441$ (see HMS for details of derivations). Here, the estimator of $\widehat{\sigma}^2_\epsilon$ is truncated at zero but our experience with Monte Carlo experiments suggests this will rarely be needed.[12]

[12] Without truncation, negative estimates of σ^2_ϵ can occur when the residuals $\widehat{\epsilon}_i$ have skewness greater than the maximum skewness that is possible with the convolution of normal and half-normal random variables. Olson et al. (1980) refer to this as "type-II failure" in the context of corrected ordinary least squares. See Papadopoulos and Parmeter (2021) for discussion.

3.3 Further Details

Regardless of whether the SVKZ or HMS estimates are used, the resulting estimates of $\phi(Z_i)$ and $r_1(Z_i)$ and their derivatives with respect to elements of Z_i can be used to recover estimates of marginal rates of substitution and transformation as well as derivatives of conditional mean efficiency in terms of the original (x, y)-space. Details are given in the separate Appendix A.

Several extensions of the ideas developed above are possible. To conserve space, we give brief discussions of these in the separate Appendix D, with systematic analysis of the ideas left for future work. In Section D.1, we show how the methods we have described above can be adapted to estimate radial input, output and hyperbolic distances. Section D.2, we discuss how exogenous environmental variables can be introduced into the analysis. This allows investigation of the effect of such variables on eh production frontier Ψ^∂ as well as on the distribution of inefficiency. Section D.3 discusses how the SVKZ and HMS estimators can be extended to impose free disposability or convexity on the production set Ψ, thereby providing stochastic analogs of FDH or DEA estimators. Finally, Section D.4 gives an outline of how the direction vector d can be allowed to be determined by data, although at a cost of increased computational burden.

4 Efficiency of US Banks

We use the methods developed above to estimate the technical efficiency of US commercial banks operating in 2003, 2008, 2013 and 2019 and thereby to examine how bank performance was affected by, and responded to, the financial crisis of 2008. We use fourth-quarter data from the US Federal Deposit Insurance Corporation Statements of Income and Condition (Call Reports). Following prior work (e.g., Wheelock and Wilson, 2018), and using subscript i to index observations $1, \ldots, n$, we define two quasi-fixed inputs, i.e., non-performing loans (X_{i1}) and equity (X_{i2}), as well as three variable inputs including purchased funds (X_{i3}), labor

(X_{i4}) and physical capital (X_{i5}). The amount of non-performing loans provides a measure or risk, and equity is a source of financial capital that is not easily adjusted in the short-run. We define three variable outputs, i.e., total loans (Y_{i1}), securities (Y_{i2}) and off-balance sheet activities (Y_{i3}) measured by total non-interest income. For 2003, 2008, 2013 and 2019 we have 7,441, 6,609, 5,468 and 4,335 observations (respectively) for a total of $n = 23,853$ observations.[13] Summary statistics for these variables are given in Table F.17 appearing in the separate Appendix F. Table F.17 indicates that marginal distributions of the input and output variables are skewed to the right, reflecting the skewness of the distribution of bank sizes measured by total assets.

As others (e.g., Wheelock and Wilson, 2001, 2012 and 2018) find, there is substantial collinearity among the specified inputs and outputs. We exploit this by reducing the dimensionality of our problem along the lines of Daraio and Simar (2007) and Wilson (2018). First, we standardize each of our eight variables by dividing by their standard deviations (this amounts to merely changing units of measurement). For the $n \times 3$ matrix $\boldsymbol{Y}$ of outputs, the $n \times 2$ matrix $\boldsymbol{X}_1$ of inactive inputs (non-performing loans and equity) and the $n \times 3$ matrix $\boldsymbol{X}_2$ of active outputs (purchased funds, labor and physical capital), we examine eigenvalues of the moment matrices $\boldsymbol{Y}'\boldsymbol{Y}$, $\boldsymbol{X}_1'\boldsymbol{X}_1$ and $\boldsymbol{X}_2'\boldsymbol{X}$. The corresponding ratios of the largest eigenvalues to the sums of the eigenvalues are 0.9487, 0.9290 and 0.9542 (respectively). The simulation results of Wilson (2018) suggest that in the absence of noise, DEA and FDH estimates of efficiency are likely to have smaller mean-square error using reduced-dimensional data as opposed to the full, eight-dimensional data in our problem. Consequently, we reduce the dimensionality to $p = 2$, $q = 1$ by using the first principle components of $\boldsymbol{X}_1$, $\boldsymbol{X}_2$ and $\boldsymbol{Y}$ obtained by multiplying these matrices by the eigenvectors corresponding to the largest eigenvalues of each of the moment matrices to obtain X_{i1}^*, X_{i2}^* and Y_i^* (respectively). Summary statistics for these variables are also given in Table F.17. We set the first element of

[13]The declining numbers of observations over time reflects on-going consolidation in the U.S. banking industry through mergers and acquisitions.

the direction vector to zero in order to hold the inactive inputs constant, and we set the second and third elements equal to the medians of X^*_{i2} and Y^*_i, ensuring that our results are independent of units of measurement. We construct the $(r \times r)$ rotation matrix R_d (where $r = 3$ after reducing dimensionality) as described in Section 2.3, and use this to translate each observation (X_i, Y_i) to (Z_i, U_i), where Z_i is $1 \times (r-1)$ and U_i is scalar. Summary statistics for the transformed data in (z, u)-space are given in Table F.17 appearing in the separate Appendix F.

In addition to Z_i, we include in our regressions both time and firm effects. For estimation of $r_1(\cdot)$ in (3.1), we regress U_i on $(Z_i,\ T_i,\ L_i)$, $i = 1,\ \ldots,\ n$, where $T_i = 1,\ 2,\ 3$ or 4 corresponding to 2003, 2008, 2013 or 2019, and L_i is a unique numeric label for the firm represented by observation i. Local linear regression requires an $(n \times n)$ diagonal matrix of kernel weights for each observation $i = 1,\ \ldots,\ n$. we use a Gaussian product kernel for the two continuous elements of Z_i. For estimating $r_1(\cdot)$ at observation i, the jth diagonal element of our weighting matrix is given by

$$\psi_j = \left[\prod_{k=1}^{2} h_k^{-1} K\left(\frac{Z_{jk} - Z_{ik}}{h_k}\right)\right] h_3^{|T_j - T_i|} h_4^{I(L_j \neq L_i)} \tag{4.1}$$

Z_{ik} denotes the kth element of Z_i, $K(\cdot)$ is the standard normal density function, $I(\cdot)$ denotes the indicator function and h_ℓ, $\ell \in \{1,\ \ldots,\ 4\}$ are bandwidths with $h_\ell \in [0, 1]$ for $\ell \in \{3,\ 4\}$. The rate of convergence in the nonparametric regressions depends on the number of continuous right-hand side variables (three in our case), and is not affected by the discrete time and firm effects (see Li and Racine, 2007 or Henderson and Parmeter, 2015 for discussion).

Bandwidths are optimized via leave-one-out cross-validation. To avoid possible numerical difficulties, we constrain the first two bandwidths (corresponding to elements of Z_i) by imposing a lower bound $0.1\left(\widehat{\sigma}_k n^{-1/(4+r^*)}\right)$ on h_k, $k \in \{1,\ 2\}$ where $\widehat{\sigma}_k$ is the sample standard deviation of Z_{ik}, $i = 1,\ \ldots,\ n$ the kth principle component and $r^* = 2$ is the number of continuous variables used in each regression. In addition, we impose an upper bound on each

h_k given by $3\,[\max_i(Z_{ik}) - \min_i(Z_{ik})]$. We obtain optimized bandwidths shown in Table F.18 in the Separate Appendix F with the corresponding lower and upper bounds. For each of the three estimated regressions, regression, optimized bandwidths for the two continuous Z-variables are well within their imposed bounds. For the regression of U_i on Z_i, the optimized values of both h_3 and h_4 are near zero, and consequently there is little smoothing across different years or across different firms. For the regressions involving squared and cubed residuals, the optimized values of h_3 remain small, but larger than for the first regression, but the optimized value of the firm-effect bandwidth h_4 is 1, indicating that no distinction in the local weighting is made between observations for the same or different firms.

After regressing U_i on Z_i, and the squared as well as the cubed residuals from this regression on Z_i, we construct two sets of estimates $\widehat{\sigma}_\eta(Z_i)$, $\widehat{\mu}_\eta(Z_i)$ and $\widehat{\sigma}_\epsilon(Z_i)$ for $i = 1, \ldots, n$. The first set of estimates is constructed using (3.8)–(3.10) in Section 3.1 as proposed by SVKZ. The second set of estimates include estimates $\widehat{\sigma}_\eta(Z_i)$ and $\widehat{\sigma}_\epsilon(Z_i)$ computed from (3.15)–(3.16) along the lines of HMS, as well as estimates $\widehat{\mu}_\eta(Z_i)$ computed from (3.9) but using the HMS estimate $\widehat{\sigma}_\eta(Z_i)$ computed from (3.15). In addition, we use both sets of estimates to compute estimates $\widehat{\zeta}(Z_i)$ of

$$\zeta(Z_i) = \frac{\text{VAR}(\eta_i(Z_i))}{\text{VAR}(\eta_i(Z_i)) + \text{VAR}(\epsilon_i(Z_i))} = \frac{\frac{\pi-2}{\pi}\sigma_\eta^2(Z_i)}{\frac{\pi-2}{\pi}\sigma_\eta^2(Z_i) + \sigma_\epsilon^2(Z_i)}. \tag{4.2}$$

in order to obtain some idea of how much noise is in the DGP. Clearly, $\zeta(Z_i) \in [0, 1]$, with $\zeta(Z_i) = 1$ if there is no noise, and $\zeta(Z_i)$ becoming smaller as the noise-to-signal ratio referred to in Section E increases.

Summary statistics for our estimates are given in Table 1. We include both sets for results for comparison purposes, but recall that our simulation results discussed in Section E suggest that the HMS estimates should be more accurate in terms of IMSE than the SVKZ estimates. Overall, both the median and the mean of the HMS efficiency estimates are larger than for the corresponding SVKZ estimates. Both sets of estimates indicate that inefficiency declined from 2003 to 2008, and from 2008 to 2013, but then increased from 2013 to 2019.

Figure 3 shows kernel estimates of the densities of the HMS efficiency estimates for each year, and these confirm the pattern observed in Table 1. Apparently, banks became more efficient—perhaps due to necessity—during and shortly after the great recession, but by 2019 had become less efficient, although still more efficient than in 2003.

Returning to Table 1, both sets of estimates $\zeta(Z_i)$ show evidence of noise in the DGP, although not for all observations. Not surprisingly, the SVKZ estimates are problematic, as for about 10-percent of the observations in each year, both $\widehat{\sigma}_\eta(Z_i)$ and $\widehat{\sigma}_\epsilon(Z_i)$ are zero, preventing $\widehat{\zeta}(Z_i)$ from being computed. Focusing on the HMS estimates of $\zeta(Z_i)$, the results for 2003 indicate noise in the DGP for less than 25-percent of the observations, and for less than 50-percent of the observations in 2019. The results also show evidence of considerably more noise in 2013 than in the other years. This was likely a period of transition for banks, falling near the end of the great recession, but with unusually low interest rates still being maintained by the Federal Reserve System. In any case, the evidence in Table 1 of noise in the DGP for banks is important. It casts doubt on the large number of studies that have used deterministic FDH or DEA estimators to examine bank efficiency, as these estimators are inconsistent in the presence of noise.[14]

5 Summary and Conclusions

The methods proposed in this paper provide a very flexible, almost fully-nonparametric approach for evaluating the performance of firms in terms of their efficiency. For purposes of estimating the production frontier, we need only some mild assumptions on continuity and smoothness of the frontier, symmetry of the density of the noise process, and tail behavior of the noise and inefficiency distributions. No functional form assumptions are needed—not for the frontier, nor for the distributions of noise and inefficiency. In addition, our method

[14] A search on Google Scholar on 29 January 2021 using the keywords "bank," "efficiency" and "data envelopment analysis" yielded about 74,200 results; a similar search replacing "data envelopment analysis" with "free disposal hull" yielded about 27,700 results.

accommodates both multiple inputs as well as multiple outputs while avoiding endogeneity problems. As discussed in Section 1, Other flexible methods that have been proposed to date either do not allow for noise, do not permit multiple inputs and outputs, involve significant computational difficulties, or involve endogeneity problems. Of course, as with most non-parametric methods, ours incurs the curse of dimensionality, but as illustrated in Section 4, dimension reduction methods can mitigate this curse in many cases.

Our empirical application on U.S. commercial banks demonstrates that our method is computationally tractable with real data. Our estimates provide evidence that, at least in our sample, the DGP for banks involves both noise and inefficiency. A large number of banking studies have used deterministic, nonparametric DEA or FDH estimators which are inconsistent when noise is present, and consequently our estimates call into question studies that have used DEA or FDH estimators.

There have been for many years two disparate strands in the literature on frontier efficiency estimation. One strand consists of papers describing deterministic estimators such as DEA and FDH estimators that require no functional form assumptions, but which do not allow for noise. The other strand consists of papers that are often fully-parametric, or in recent years, partially parametric, that permit noise, but do not allow multiple inputs and outputs. We hope that our paper provides a bridge between these two worlds.

References

Aigner, D., C. A. K. Lovell, and P. Schmidt (1977), Formulation and estimation of stochastic frontier production function models, *Journal of Econometrics* 6, 21–37.

Atkinson, S. E. and M. G. Tsionas (2016), Directional distance functions: Optimal endogenous directions, *Journal of Econometrics* 190, 301–314.

Bahadur, R. R. and L. J. Savage (1956), The nonexistence of certain statistical procedures in nonparametric problems, *The Annals of Mathematical Statistics* 27, 1115–1122.

Battese, G. E. and G. S. Corra (1977), Estimation of a production frontier model: With application to the pastoral zone off eastern australia, *Australian Journal of Agricultural Economics* 21, 169–179.

Bauer, P. W. (1990), Recent developments in the econometric estimation of frontiers, *Journal of Econometrics* 46, 39–56.

Bravo-Ureta, B. E. and A. E. Pinheiro (1993), Efficiency analysis of developing country agriculture: A review of the frontier function literature, *Agricultural and Resource Economics Review* 22, 88–101.

Chambers, R. G., Y. Chung, and R. Färe (1996), Benefit and distance functions, *Journal of Economic Theory* 70, 407–419.

— (1998), Profit, directional distance functions, and nerlovian efficiency, *Journal of Optimization Theory and Applications* 98, 351–364.

Coelli, T. (1995), Estimators and hypothesis tests for a stochastic frontier function: A monte carlo analysis, *Journal of Productivity Analysis* 6, 247–268.

Coelli, T. and S. Perelman (1996), Efficiency measurement, multiple-output technologies and distance functions: With application to European railways. CREPP working paper 96/05, University of Liège, Belgium.

Cuesta, R. A. and L. Orea (2002), Mergers and technical efficiency in Spanish savings banks: A stochastic distance function approach, *Journal of Banking and Finance* 26, 2231––2247.

Daouia, A., J. P. Florens, and L. Simar (2020), Robust frontier estimation from noisy data: A tikhonov regularization approach, *Econometrics and Statistics* 14, 1–23.

Daraio, C. and L. Simar (2007), *Advanced Robust and Nonparametric Methods in Efficiency Analysis*, New York: Springer Science+Business Media, LLC.

Fan, J. and I. Gijbels (1996), *Local Polynomial Modelling and Its Applications*, London: Chapman and Hall.

Fan, Y., Q. Li, and A. Weersink (1996), Semiparametric estimation of stochastic production frontier models, *Journal of Business and Economic Statistics* 14, 460–468.

Färe, R., S. Grosskopf, and D. Margaritis (2008), Productivity and efficiency: Malmquist and more, in H. Fried, C. A. K. Lovell, and S. Schmidt, eds., *The Measurement of Productive Efficiency*, chapter 5, Oxford: Oxford University Press, 2nd edition, pp. 522–621.

Färe, R., C. Pasurka, and M. Vardanyan (2017), On endogenizing direction vectors in parametric directional distance function-based models, *European Journal of Operational Research* 262, 361–369.

Florens, J. P., L. Simar, and I. Van Keilegom (2020), Estimation of the boundary of a variable observed with symmetric error, *Journal of the American Statistical Association* 115, 425–441.

Hafner, C. M., H. Manner, and L. Simar (2018), The "wrong skewness" problems in stochastic frontier models: A new approach, *Econometric Reviews* 37, 380–400.

Hall, P. and L. Simar (2002), Estimating a changepoint, boundary or frontier in the presence of observation error, *Journal of the American Statistical Association* 97, 523–534.

Henderson, D. J. and C. F. Parmeter (2015), *Applied Nonparametric Econometrics*, New York: Cambridge University Press.

Jondrow, J., C. A. K. Lovell, I. S. Materov, and P. Schmidt (1982), On the estimation of technical inefficiency in the stochastic frontier production model, *Journal of Econometrics* 19, 233–238.

Kneip, A., L. Simar, and I. Van Keilegom (2015), Frontier estimation in the presence of measurement error with unknown variance, *Journal of Econometrics* 184, 379–393.

Kneip, A., L. Simar, and P. W. Wilson (2008), Asymptotics and consistent bootstraps for DEA estimators in non-parametric frontier models, *Econometric Theory* 24, 1663–1697.

Kumbhakar, S. C., B. U. Park, L. Simar, and E. G. Tsionas (2007), Nonparametric stochastic frontiers: A local likelihood approach, *Journal of Econometrics* 137, 1–27.

Kumbhakar, S. C., C. F. Parmeter, and V. Zelenyuk (2020), Stochastic frontier analysis: Foundations and advances i.

Kuosmanen, T. and A. Johnson (2017), Modeling joint production of multiple outputs in StoNED: Directional distance function approach, *European Journal of Operational Research* 262, 792—801.

Kuosmanen, T. and M. Kortelainen (2012), Stochastic non-smooth envelopment of data: Semi-parametric frontier estimation subject to shape constraints, *Journal of Productivity Analysis* 38, 11–28.

Layer, K., A. L. Johnson, R. C. Sickles, and G. D. Ferrier (2020), Direction selection in stochastic directional distance functions, *European Journal of Operational Research* 280, 351–364.

Li, Q. and J. Racine (2007), *Nonparametric Econometrics*, Princeton, NJ: Princeton University Press.

Lovell, C. A. K. (1993), Production frontiers and productive efficiency, in H. Fried, C. A. K. Lovell, and S. Schmidt, eds., *The Measurement of Productive Efficiency: Techniques and Applications*, Oxford: Oxford University Press, Inc., pp. 3–67.

Malikov, E., S. C. Kumbhakar, and M. G. Tsionas (2016), A cost system approach to the stochastic directional technology distance functions with undesirable outputs: The case of US banks in 2011–2010, *Journal of Applied Econometrics* 31, 1407–1429.

Meeusen, W. and J. van den Broeck (1977), Efficiency estimation from Cobb-Douglas production functions with composed error, *International Economic Review* 18, 435–444.

Olson, J. A., P. Schmidt, and D. M. Waldman (1980), A monte carlo study of estimators of stochastic frontier production functions, *Journal of Econometrics* 13, 67–82.

Papadopoulos, A. and C. F. Parmeter (2021), Type II failure and specification testing in the stochastic frontier model, *European Journal of Operational Research* 293, 990–1001.

Park, B. U., L. Simar, and V. Zelenyuk (2015), Categorical data in local maximum likelihood: Theory and application to productivity analysis, *Journal of Productivity Analysis* 43, 199–214.

Parmeter, C. F. and V. Zelenyuk (2019), Combining the virtues of stochastic frontier and data envelopment analysis, *Operations Research* 67, 1628–1658.

Richmond, J. (1974), Estimating the efficiency of production, *International Economic Review* 15, 515–521.

Serra, T., A. O. Lansink, and S. E. Stefanou (2011), Measurement of dynamic efficiency: A directional distance function parametric approach, *American Journal of Agricultural Economics* 93, 756–767.

Sickles, R. C. and V. Zelenyuk (2019), *Measurement of Productivity and Efficiency*, Cambridge, UK: Cambridge University Press.

Simar, L. (2007), How to improve the performances of DEA/FDH estimators in the presence of noise, *Journal of Productivity Analysis* 28, 183–201.

Simar, L., I. Van Keilegom, and V. Zelenyuk (2017), Nonparametric least squares methods for stochastic frontier models, *Journal of Productivity Analysis* 47, 189–204.

Simar, L. and P. W. Wilson (2010), Estimation and inference in cross-sectional, stochastic frontier models, *Econometric Reviews* 29, 62–98.

— (2013), Estimation and inference in nonparametric frontier models: Recent developments and perspectives, *Foundations and Trends in Econometrics* 5, 183–337.

— (2015), Statistical approaches for non-parametric frontier models: A guided tour, *International Statistical Review* 83, 77–110.

Simar, L. and V. Zelenyuk (2011), Stochastic FDH/DEA estimators for frontier analysis, *Journal of Productivity Analysis* 36, 1–20.

Vardanyan, M. and D.-W. Noh (2006), Approximating pollution abatement costs via alternative specifications of a multi-output production technology: A case of the US electric utility industry, *Journal of Environmental Management* 80, 177–190.

Wheat, P., W. Greene, and A. Smith (2014), Understanding prediction intervals for firm specific inefficiency scores from parametric stochastic frontier models, *Journal of Productivity Analysis* 42, 55–65.

Wheelock, D. C. and P. W. Wilson (2001), New evidence on returns to scale and product mix among U.S. commercial banks, *Journal of Monetary Economics* 47, 653–674.

— (2008), Non-parametric, unconditional quantile estimation for efficiency analysis with an application to Federal Reserve check processing operations, *Journal of Econometrics* 145, 209–225.

— (2012), Do large banks have lower costs? new estimates of returns to scale for U.S. banks, *Journal of Money, Credit, and Banking* 44, 171–199.

— (2018), The evolution of scale-economies in U.S. banking, *Journal of Applied Econometrics* 33, 16–28.

Wilson, P. W. (2011), Asymptotic properties of some non-parametric hyperbolic efficiency estimators, in I. Van Keilegom and P. W. Wilson, eds., *Exploring Research Frontiers in Contemporary Statistics and Econometrics*, Berlin: Springer-Verlag, pp. 115–150.

— (2018), Dimension reduction in nonparametric models of production, *European Journal of Operational Research* 267, 349–367.

Ziegelmann, F. (2002), Nonparametric estimation of volatility functions: The local exponential estimator, *Econometric Theory* 18, 985–991.

Table 1: Efficiency Estimates and Variance Ratios for U.S. Banks

Year	Min	Q1	Median	Mean	Q3	Max	NA
			—— HMS Efficiency Estimates, $\widehat{\mu}_\eta(Z)$ ——				
2003	0.0033	0.0873	0.1107	0.1179	0.1224	1.0759	—
2008	0.0017	0.0719	0.0993	0.1186	0.1291	1.6144	—
2013	0.0030	0.0493	0.0660	0.0941	0.0896	2.4535	—
2019	0.0080	0.0890	0.1051	0.1334	0.1210	1.8958	—
			—— HMS Estimates, $\widehat{\zeta}(Z)$ ——				
2003	0.0050	1.0000	1.0000	0.9271	1.0000	1.0000	0
2008	0.0035	0.9081	0.9955	0.9422	1.0000	1.0000	0
2013	0.0005	0.2670	0.5094	0.5684	1.0000	1.0000	0
2019	0.0070	0.6930	1.0000	0.8296	1.0000	1.0000	0
			—— SVKZ Efficiency Estimates, $\widehat{\mu}_\eta(Z)$ ——				
2003	0.0000	0.0533	0.0951	0.0788	0.1161	0.2393	—
2008	0.0000	0.0502	0.0854	0.0932	0.1268	0.3061	—
2013	0.0000	0.0000	0.0425	0.0370	0.0627	0.1830	—
2019	0.0000	0.0202	0.0906	0.0804	0.1071	1.4154	—
			—— SVKZ Estimates, $\widehat{\zeta}(Z)$ ——				
2003	0.0000	0.9189	1.0000	0.8453	1.0000	1.0000	965
2008	0.0000	0.8840	0.9501	0.8797	1.0000	1.0000	759
2013	0.0000	0.0000	0.2576	0.3492	0.5737	1.0000	626
2019	0.0000	0.3699	0.9269	0.6829	1.0000	1.0000	459

NOTE: "NA" refers to "not available," i.e., the number of observations where an estimate cannot be computed. This is relevant only for the estimates $\widehat{\zeta}(Z_i)$, and occurs when the denominator equals zero. As mentioned in Section 4, the numbers of observations in each year are 7,441, 6,609, 5,468 and 4,335 for 2003, 2008, 2013 and 2019 (respectively).

Figure 1: Transformation from (x, y)-Space to (z, u)-Space

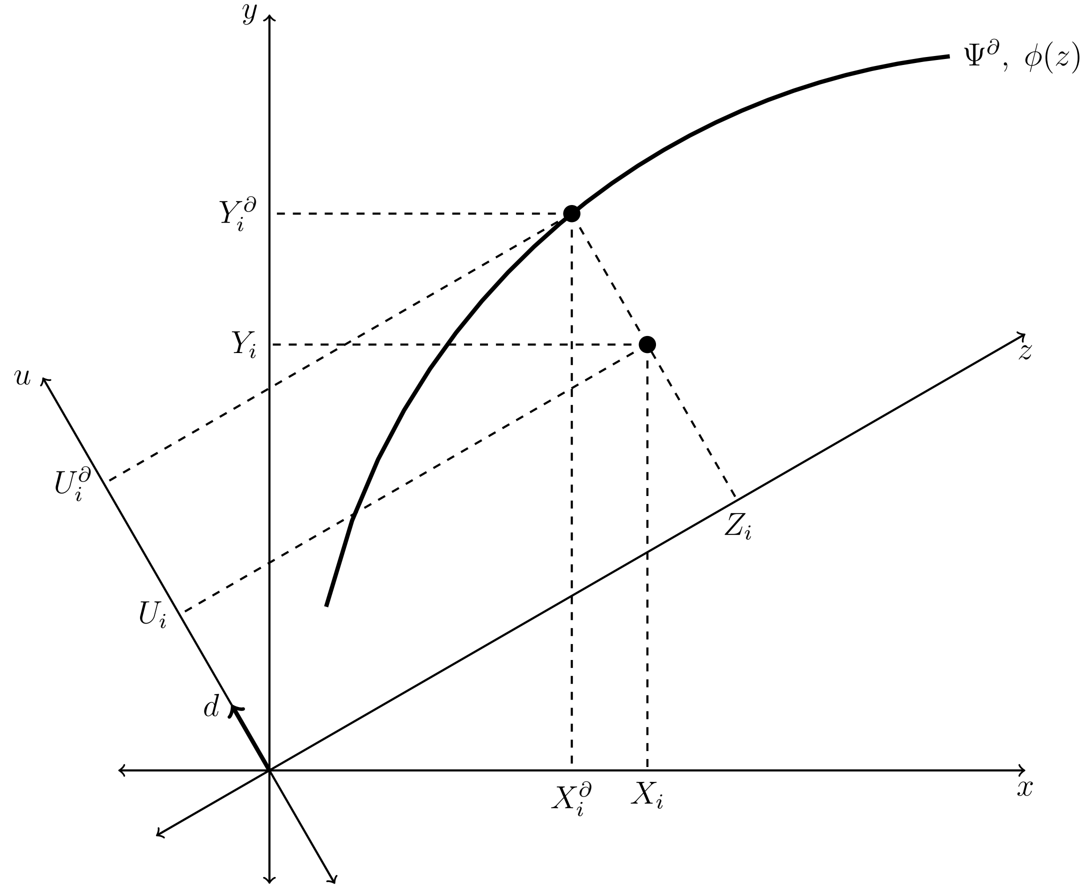


Figure 2: Half-Normal Inefficiency with Various Symmetric Noise Distributions

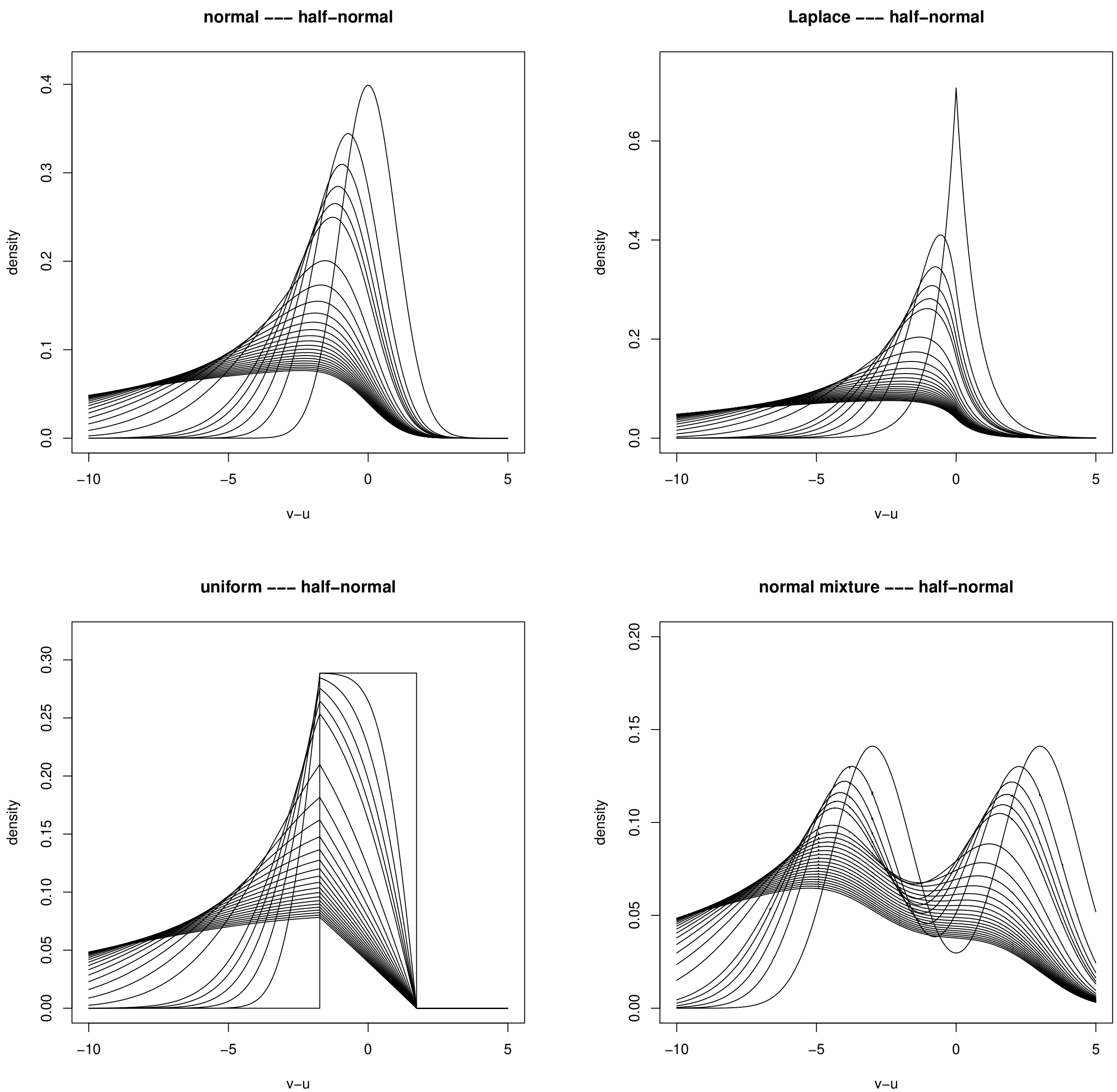
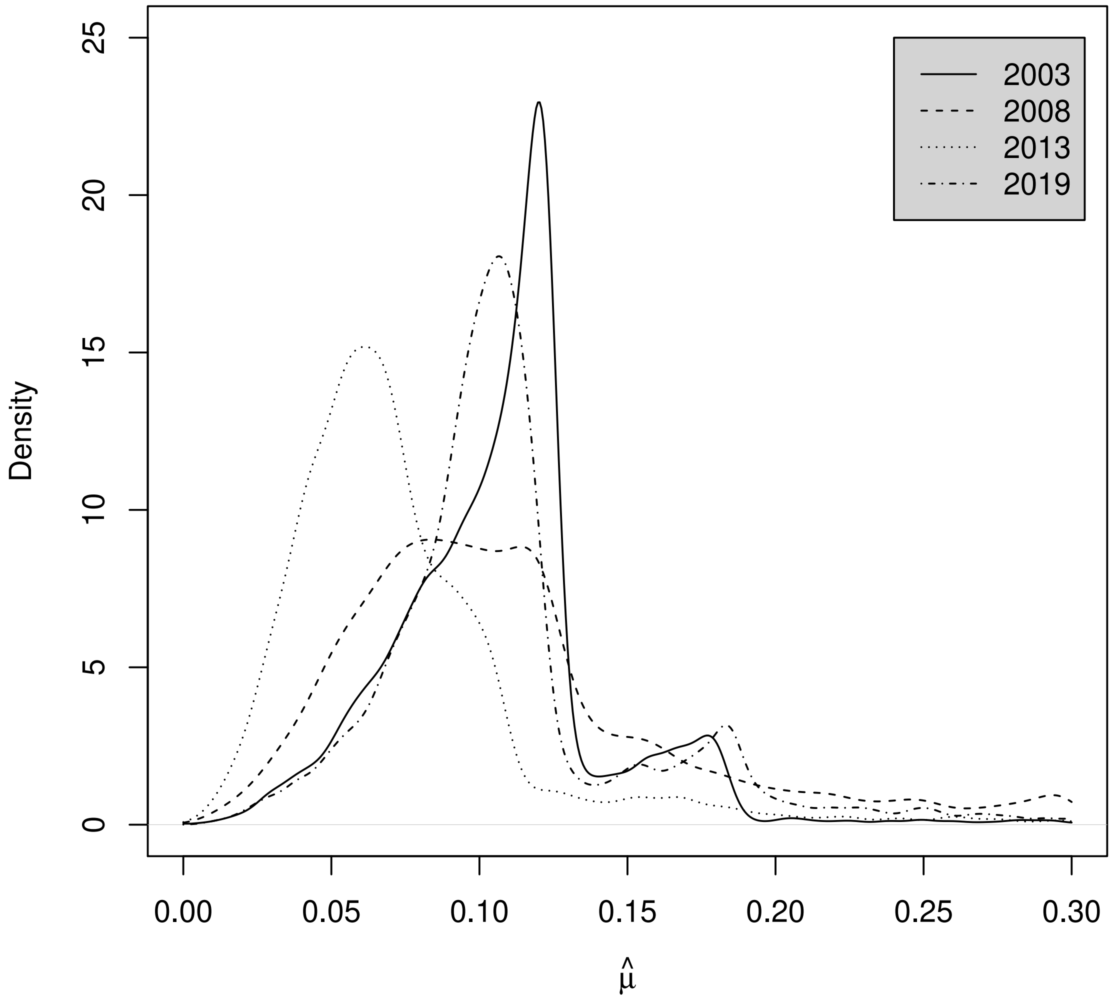


Figure 3: Kernel Estimates of Densities of US Banks' Inefficiency (HMS)



Nonparametric, Stochastic Frontier Models with Multiple Inputs and Outputs: Supplementary Material

Léopold Simar Paul W. Wilson*

August 2021

*Simar: Institut de Statistique, Biostatistique, et Sciences Actuarielles, Université Catholique de Louvain-la-Neuve, Louvain-la-Neuve, Belgium; email: leopold.simar@uclouvain.be.

Wilson: Department of Economics and School of Computing, Clemson University, Clemson, South Carolina, USA; email: pww@clemson.edu.

We are grateful for (i) the Palmetto cluster operated by the Cyber Infrastructure Technology Integration group at Clemson University, and (ii) high-performance computing equipment purchased with funds from U.S. National Science Foundation MRI award no. 1725573. Any remaining errors are solely our responsibility.

Contents

List of Tables

A Links between the Two Spaces

Recall that the direction d and the chosen matrix V_d are fixed as discussed earlier footnote 8 in Section 2.3. In applied economic studies, researchers are often interested in marginal effects such as marginal productivity, marginal rates of substitution, etc. As demonstrated below, it is possible to recover partial derivatives in the original (x, y)-space from derivatives obtained in the new, transformed (z, u)-space.

First, the locus of the frontier function in the transformed model is given by the set of points $(z, \phi(z))$, where $z \in \mathbb{R}^{r-1}$. From this the frontier surface in (x, y) space (an $(r-1)$-manifold) can be obtained from

$$w^\partial = \begin{bmatrix} x^\partial \\ y^\partial \end{bmatrix} = R'_d \begin{bmatrix} z \\ \phi(z) \end{bmatrix} = \begin{bmatrix} V_d & \frac{d}{||d||} \end{bmatrix} \begin{bmatrix} z \\ \phi(z) \end{bmatrix}, \tag{A.1}$$

which can be viewed as a mapping $F\colon \mathbb{R}^{r-1} \mapsto \mathbb{R}^r$, where $w^\partial = F(z)$ is a manifold *parameterized* by F. The Jacobian of the transformation in (A.1) can be written as the $r \times (r-1)$ matrix

$$J_\phi(z) = \left\{ \frac{\partial w_i^\partial}{\partial z_j} \right\} = V_d + \frac{d}{||d||} \frac{\partial \phi(z)}{\partial z'}, \tag{A.2}$$

where $i = 1, \ldots, r$ and $j = 1, \ldots, (r-1)$. Existence of the Jacobian in (A.2) is ensured by Assumption 2.6.

Now consider a fixed point $(z_0, \phi(z_0))$ on the frontier. Then from (A.1) the image on the frontier manifold in w-space is $w_0^\partial = F(z_0)$. Standard results from real analysis (e.g., Lang, 1964, pp. 164–174) ensure that under Assumption 2.6,

$$F(z_0 + h) = F(z_0) + J_\phi(z_0)h + ||h||\, G(h), \tag{A.3}$$

where h is a small vector in $\mathbb{R}^{r-1}$ and $\lim_{||h|| \to 0} G(h) = 0$. Therefore, $J_\phi(z_0)$ defines a linear map from $\mathbb{R}^{r-1}$ to $\mathbb{R}^r$ which can be viewed as an approximation of F near the point z_0, up to an error term whose magnitude is small. This linear map is said to be a tangent hyperplane to F at the point z_0. Since F is differentiable at z_0 by Assumption 2.6, this hyperplane is unique due to Theorem 1 of Lang (1964, p. 166).

Clearly, $w = F(z_0 + h)$ in (x, y)-space. Therefore, the equation of the unique hyperplane tangent to the surface $F(z)$ at the point z_0 is given by

$$w = w_0^\partial + J_\phi(z_0)h. \tag{A.4}$$

This hyperplane can be written in explicit form by the linear equation

$$c'\big(w - w_0^\partial\big) = c'J_\phi(z_0)h = 0 \tag{A.5}$$

where $c \in \mathbb{R}^r$. The vector c is identified up to a multiplicative constant by the equation $c'J_\phi(z_0) = 0'_{r-1}$, or equivalently,

$$J'_\phi(z_0)c = 0_{r-1}. \tag{A.6}$$

In other words, c is determined (up to a scalar constant) by a basis of the null space of the matrix $J'_\phi(z_0)$.

Using the equation of the tangent hyperplane in (A.5), it is obvious that any component w_ℓ of w can be expressed as a function of the other elements w_k, $k = 1, \ldots, r$, $k \neq \ell$. Hence partial derivatives in (x, y)-space can be written as

$$\left\{\frac{\partial w_\ell}{\partial w_k}\right\}_{w=w_0^\partial} = -\frac{c_k}{c_\ell} \tag{A.7}$$

provided $c_\ell \neq 0$. As shown below in Section 3, the result in (A.7) can be used to estimate marginal products $\partial Y_\ell/\partial X_k$ $\forall$ $k \in \{1, \ldots, p\}$ and $\ell \in \{1, \ldots, q\}$, marginal rates of substitution $\partial X_\ell/\partial X_k$ $\forall$ $k, \ell \in \{1, \ldots, p\}$, $k \neq \ell$, and marginal rates of transformation $\partial Y_\ell/\partial Y_k$ $\forall$ $k, \ell \in \{1, \ldots, q\}$, $k \neq \ell$. If $c_\ell = 0$, then the derivative in (A.7) is undefined. However, in this case,

$$\left\{\frac{\partial w_k}{\partial w_\ell}\right\}_{w=w_0^\partial} = 0 \tag{A.8}$$

provided $c_k \neq 0$, indicating that the variable w_ℓ has no effect on the component w_k at the particular frontier point of interest.

Another interesting feature of the transformed model is the function $\mu_\eta(z)$ defined in (2.22). It is clear that the vector $\partial\mu_\eta(w^\partial)/\partial w^\partial$ can be recovered from knowledge of the vector $\partial\mu_\eta(z)/\partial z$. Since $z = V_d' w^\partial$, the Jacobian of the mapping from w^∂ to z is

$$J_\mu = \left[\frac{\partial z_i}{\partial w_j^\partial}\right] = V_d', \tag{A.9}$$

with $i = 1, \ldots, r-1$ and $j = 1, \ldots, r$. Applying the chain rule yields

$$\frac{\partial\mu_\eta(w^\partial)}{\partial w^\partial} = V_d \frac{\partial\mu_\eta(z)}{\partial z}, \tag{A.10}$$

which is an r-vector of derivatives of mean inefficiency with respect to the components of w at a point w^∂.

As shown in Section 3, the derivatives $\partial\mu_\eta(z)/\partial z$ and $\partial\phi(z)/\partial z$ can be estimated easily using local-linear (or local-quadratic) estimation of the transformed model. Let $\widehat{\phi}(Z_i)$ and $\widehat{r}_1(Z_i)$ denote estimates of the frontier function $\phi(Z_i)$ and the conditional mean function $r_1(Z_i)$ defined in Section 3.1 of the main paper. Then by inverting the transformation in (2.11), the points $(Z_i, \widehat{\phi}(Z_i))$ can be translated to estimates of the frontier surface in the original (x, y) coordinate system. This leads to

$$\widehat{W}_i^\partial = \begin{bmatrix} X_i^\partial \\ Y_i^\partial \end{bmatrix} = R_d' \begin{bmatrix} Z_i \\ \widehat{\phi}(Z_i) \end{bmatrix}, \tag{A.11}$$

where the matrix R_d defined earlier is known and non-stochastic. Similarly, estimates of points on the average production function $r_1(z)$ can be obtained in the original space via

$$\widehat{W}_i^{\text{aver}} = \begin{bmatrix} X_i^{\text{aver}} \\ Y_i^{\text{aver}} \end{bmatrix} = R_d' \begin{bmatrix} Z_i \\ \widehat{r}_1(Z_i) \end{bmatrix}. \tag{A.12}$$

Since these transformations are nonstochastic, it is clear that the asymptotic properties derived by SVKZ remain valid in the original space with the same rate of convergence (i.e. $n^{2/(p+q+3)}$).

The same is true for the other quantities of interest derived in Section A. In particular, the derivatives of efficiency with respect to inputs and outputs given in (A.10) can be estimated by noting first that differentiation in (3.5) yields

$$\frac{\partial \mu_\eta(z)}{\partial z} = \sqrt{\frac{2}{\pi}} \frac{\partial \sigma_\eta(z)}{\partial z}. \tag{A.13}$$

The derivative $\frac{\partial \sigma_\eta(z)}{\partial z}$ can be estimated by differentiating either the SVKZ estimator of $\sigma_\eta(z)$ in (3.8) or the HMS estimator in (3.15). Either of these depend on $\frac{\widehat{r}_3(z)}{\partial z}$, estimates of which are provided by the LLLS estimation of the regression of the cubed residuals $\widehat{\xi}_i^3$ on the Z_i, $i = 1, \ldots, n$. Repeated substitution then leads to estimates of the derivative on the right-hand side of (A.10), and hence estimates of $\partial \mu_\eta(w^\partial)/\partial w^\partial$.

Estimates of the marginal effects given by (A.7) can also be computed, giving estimates of marginal products, marginal rates of substitution and marginal rates of transformation. LLLS estimation of the average production function $r_1(z)$ provides estimates of $\partial r_1(Z_i)/\partial Z_i$, and from the discussion above we have estimates of $\partial \mu_\eta(Z_i)/\partial Z_i$. Adding these give estimates of $\partial \phi(Z_i)/\partial Z_i$ due to (3.11), and these can be used to compute estimates $\widehat{J}_\phi(Z_i)$ of the Jacobian matrix $J_\phi(Z_i)$ given in A.2. Substituting $\widehat{J}_\phi(Z_i)$ for $J_\phi(Z_i)$ in (A.6) and solving for the vector c allows estimates of the marginal effects given in (A.7) to be computed.

B Some Alternative Approaches for Estimation

As discussed in Section 3 of the main paper, estimation by methods other than those of SVKZ and HMS could be used to estimate our transformed model. For example, one could specify the link function as a translog function of the elements of Z_i, and further specify the distribution of $\epsilon_i \mid Z_i$ as normal and the distribution of $\eta_i \mid Z_i$ as half-normal. If, in addition, specific functional forms are specified for $\sigma_\epsilon(Z_i)$ and $\sigma_\eta(Z_i)$, the resulting fully parametric model could then be estimated by maximum likelihood. However, one or more of the required specification assumptions are likely to be wrong.[1]

[1] While maximum likelihood estimators achieve root-n convergence in fully parametric models, this only ensures rapid convergence to something that is not a feature of the true model if the model is mis-specified. Robinson (1988) refers to this as "root-n *in*consistency."

The goal here is to avoid restrictive parametric assumptions, but this raises questions regarding identification in the sense that we wish to be able to use knowledge about the conditional density $f(u \mid z)$ of U to recover information about triple $\{\phi(z),\ f(\epsilon|z),\ f(\eta|z)\}$ with ϵ and η being independent conditionally on z. For a given z, this amounts to a deconvolution problem since $u = \phi(z) + \epsilon - \eta$. So far, the discussions about identification has been in terms of the univariate dependent (output) variable case. But now, due to the change of coordinates, we have the transformed model (2.17)–(2.19) which also has a univariate dependent variable, namely U.

The results of Hall and Simar (2002) indicate that if the density of $\eta \mid z$ is completely unspecified and the density of $\epsilon \mid z$ is symmetric around zero, the model is not identified. For this reason, Hall and Simar (2002) consider a limiting case where the conditional variance of the noise $\sigma_\epsilon^2(z)$ converges to zero as the sample size n increases. Under this assumption, Hall and Simar derive consistent estimators of the model. Of course this technical assumption is rather restrictive, limiting its interest for practitioners.

Kumbhakar et al. (2007) propose a different approach, wherein the conditional densities of ϵ and η are specified *locally*, e.g., $\epsilon \mid z \sim N(0, \sigma_\epsilon^2(z))$ and $\eta \mid z \sim N^+(0, \sigma_\eta^2(z))$. The model is then clearly identified for each z, and the model can be estimated by local maximum likelihood. To maintain the nonparametric structure of the model, both scale functions are functional, and local polynomial approximations are used. Although the approach is appealing, estimation involves a formidable computational burden. Moreover, the local structure of the stochastic components is parametric, and asymptotic properties of the estimators rely on the specific parametric assumptions used to derive the local likelihood function.

Kneip et al. (2015) leave the inefficiency distribution unspecified, but require the noise $\epsilon \mid z$ to be Gaussian with unknown variance. Under the latter assumption, they prove the model is identified. For estimation, they use quasi-likelihood approaches with histogram estimators for the density of the inefficiency $\eta \mid z$. Here again the numerical burden is significant.

Daouia et al. (2020) use an alternative approach based on nonparametric estimation of the conditional survival function of $U \mid Z$, but for identification must assume that the noise follows a fully known distribution, e.g., $\epsilon \mid z \sim N(0, \sigma_\epsilon^2(z))$ where $\sigma_\epsilon^2(z)$ is known.

More recently, Florens et al. (2020) specify a model where the noise is only assumed to be symmetric (as in our specification) and a flexible parametric density is chosen for the inefficiency. They also investigate identification issues (see their Section 2) and prove that if the density of $\eta \mid z$ is completely specified by its odd cumulants of order larger than 3, then the model is identified. This is a sufficient condition for identification when the noise is symmetric.

C Alternative Approaches for Dealing with Unexpected Skewness

C.1 Sign-Constrained Local Linear Least Squares

The method is briefly mentioned in Remark 2.2 in SVKZ, where the authors suggest first computing the ordinary, local-linear least squares (LLLS) nonparametric estimator $\widehat{r}_3(z)$ of $r_3(z)$. If the resulting estimate has the wrong sign (i.e., if $\widehat{r}_3(z) > 0$), then a new, constrained LLLS estimate can be computed while imposing the constraint $r_3(z) \leq 0$ in a constrained optimization problem. We can use the same idea here, but we show that the constrained optimization problem can be easily solved by adapting results from Liew (1976) derived for parametric linear least squares.

Consider the LLLS estimator of the mean regression function $m(z) = \mathbb{E}(\zeta | Z = z)$ for a univariate response ζ and where z has dimension $(r-1)$. Here, $\zeta_i = \widehat{\varepsilon}_i^3$ for estimating $r_3(z)$, but it could also be $\widehat{\varepsilon}_i^2$ for estimating $r_2(z)$.

It is well known (see Fan and Gijbels, 1996, Section 7.8) that LLLS estimator can be written as

$$\widehat{\beta} = (\mathcal{Z}_D' \mathcal{W} \mathcal{Z}_D)^{-1} \mathcal{Z}_D' \mathcal{W} \mathcal{Y}, \tag{C.1}$$

where $\mathcal{Y} = (\zeta_1, \ldots, \zeta_n)'$, $\mathcal{Z}_D = \left[i_n \quad (\mathcal{Z} - i_n z') \right]$ where $\mathcal{Z} = (Z_1, \ \ldots, \ Z_n)'$ and the usual weights matrix is $\mathcal{W} = \text{diag}(K_h(Z_i - z))$. The LLLS estimator

$$\widehat{\beta} = \arg\min_{\beta} (\mathcal{Y} - \mathcal{Z}_D \beta)' \mathcal{W} (\mathcal{Y} - \mathcal{Z}_D \beta) \tag{C.2}$$

is the solution of a weighted least squares problem, and

$$\widehat{m}(z) = \widehat{\beta}_1 \tag{C.3}$$

and

$$\widehat{\frac{\partial m}{\partial z_j}}(z) = \widehat{\beta}_{j+1} \tag{C.4}$$

for $j = 1, \ \ldots, \ r-1$.

Following the arguments in Liew (1976), it can be shown that the solution of the weighted least squares problem under the constraint that $\beta_1 \leq 0$, when the constraint is binding, is given by the simple expression

$$\widehat{\beta}_C = \widehat{\beta} - A_1 \widehat{\beta}_1 / a_1, \tag{C.5}$$

where A_1 is the first column of the matrix $(\mathcal{Z}_D' \mathcal{W} \mathcal{Z}_D)^{-1}$ and a_1 is the first element of A_1. It can be shown that the same correction is valid for the case of the constraint $\beta_1 \geq 0$, when the latter is binding.

C.2 Local Exponential smoothing

To be complete in our presentation, we briefly summarize the steps for obtaining the Local Exponential estimator proposed by Ziegelmann (2002) for regression models where the dependent variable is sign constrained (say positive). In our case (for the 3rd moment) it is obtained by solving

$$(\hat{\alpha}, \hat{\beta}) = \arg\min_{\alpha, \beta} \sum_{i=1}^{n} \left\{ \left[-\widehat{\varepsilon}_i^3 \right] - \exp\left[\alpha + \beta'(Z_i - z) \right] \right\}^2 K_h(Z_i - z) \tag{C.6}$$

for a given z. Then we have $\widehat{r}_3(z) = -\exp(\hat{\alpha}) \leq 0$. Ziegelmann shows that this estimator asymptotic properties similar to those of the usual local linear estimator. However, the numerical burden of the local exponential estimator is large, and numerical instabilities may arise due to the exponential transformation in (C.6). In practice, to eliminate the risk of ending with local minima, several starting values should be tried. We conducted some initial Monte Carlo experiments using this estimator, but even using robust nonlinear optimization techniques, e.g., the Nelder and Mead (1965) simplex method, the results indicate a high degree of numerical instability, frequently yielding unreliable results and in all cases incurring large computational cost.

D Some Extensions

D.1 The Cases of Hyperbolic and Radial Distances

The directional distance function defined in (2.3) is *additive* in the sense that for a given point $(x, y) \in \Psi$, subtracting $\delta(x, y \mid d)d_x$ from x and adding $\delta(x, y \mid d)d_y$ to y projects (x, y) onto Ψ^∂ in the direction $d = (-d_x', d_y)$. Alternatively, one might be interested in the Farrell hyperbolic measure of efficiency

$$\gamma(x, y) := \inf\left\{\gamma \mid (\gamma x, \gamma^{-1}y) \in \Psi\right\} \tag{D.1}$$

described by Färe et al. (1985). The methods described above can easily be adapted to estimate such measures by working with logs of (X_i, Y_i) provided all of the input-output data are strictly positive.. Analogous to (2.4), we can write

$$\begin{bmatrix}\log X_i \\ \log Y_i\end{bmatrix} = \begin{bmatrix}\log X_i^\partial \\ \log Y_i^\partial\end{bmatrix} + \begin{bmatrix}-(\epsilon_i - \eta_i)i_p \\ (\epsilon_i - \eta_i)i_q\end{bmatrix} \tag{D.2}$$

where ϵ_i and η_i are as in Assumption 2.4. In the absence of noise, we have $\eta_i = \log\gamma(X_i, Y_i) \geq 0$, but noise can be accommodated exactly as before.

Alternatively, one might consider the Farrell (1957) output measure of efficiency given by

$$\lambda(x, y) := \sup \{\lambda \mid (x, \lambda y) \in \Psi\}. \tag{D.3}$$

Estimates can be obtained similarly using $d = (0_p', i_q')'$ and logged data. In the absence of noise, we have $\eta_i = \log \lambda(x, y)$, but noise can be allowed for as before. Yet another alternative is the Farrell (1957) input measure given by

$$\theta(x, y) := \inf \{\theta \mid (\theta x, y) \in \Psi\}. \tag{D.4}$$

This can be estimated by setting $d = (-i_p', 0_q')'$ and again using logged data, leading to $\eta_i = \log \theta(x, y)$ in the absence of noise.

D.2 Introducing Exogenous Environmental Variables

The models developed in Sections 2.2 and 2.3 can be easily extended to accommodate environmental variables E_i that may influence either the frontier levels W_i^∂ or characteristics of the components of ξ_i. Conditioning the frontier points W_i^∂ on the E_i leads to $W_i^\partial(E_i) = (X_i^\partial(E_i), Y_i^\partial(E_i))$, and as a consequence the scale functions $\widetilde{\sigma}_\epsilon$ and $\widetilde{\sigma}_\eta$ must also be conditioned on the E_i. This amounts to replacing (2.17)–(2.19) with the conditional transformed model

$$U_i = \phi(Z_i, E_i) + \epsilon_i - \eta_i \tag{D.5}$$

where $\epsilon_i \mid Z_i, E_i \sim \text{Sym}(0, \sigma_\epsilon^2(Z_i, E_i))$ and $\eta_i \mid Z_i, E_i \sim D_+(\sigma_\eta(Z_i, E_i))$.

This is sufficient to allow one to disentangle to role of E on the frontier level as well as on the inefficiency distribution, thereby avoiding restrictive assumptions that are often made in in the case of no noise (see Simar and Wilson, 2007 for discussion). For estimation, one can condition everywhere on not just Z_i but instead (Z_i, E_i) and follow the route outlined above in Section 3. All of the results remain valid with the exception of the rate of convergence, which is slower when environmental variables are present. For $E \in \mathbb{R}^s$, the optimal rate is $n^{2/(p+q+s+3)}$, which makes clear that including environmental variables exacerbates the curse of dimensionality, leading to perhaps imprecise estimates unless n is large.

D.3 Stochastic Versions of FDH/DEA-Type Estimators

Both the SVKZ and HMS estimators discussed in Section 3 can be extended to impose assumptions of free disposability or convexity of the production set Ψ. It is sufficient to apply the FDH estimator (to impose free disposability) or the DEA estimator (to impose both free disposability and convexity) to the estimated frontier points $\widehat{W}_i^{\partial}$ obtained with either the SVKZ or HMS estimator as discussed above in Section 3. This amounts to a regularization step, as discussed Simar and Zelenyuk (2011). The usual convergence rates of the FDH and DEA estimators are $n^{1/(p+q)}$ and $n^{2/(p+q+1)}$), respectively, but here the rates can be improved.

To achieve this, frontier points on a grid of values $\{z_k\}_{k=1}^{K_n}$ should be estimated using the methods described in Section 3, with K_n large enough to control the size of the error of the regularization. For instance, if the FDH estimator is applied on these K_n points, the rate of error in the FDH regularization step is $K_n^{1/(p+q)}$. Hence, in order to obtain a free-disposal estimator of Ψ keeping the same properties as the initial, un-regularized estimator, one should select $K_n > n^{2(p+q)/(p+q+3)}$. Alternatively, if the DEA estimator is used, one should choose $K_n > n^{(p+q+1)/(p+q+3)}$ for the regularization step.

D.4 Data-Driven Direction

As discussed in footnote 4 of Section 2, some papers allow the direction vector d to be determined by data according to some criterion function which is then optimized. Often, the criterion involves minimizing inefficiency across firms in the sample. If we let d be constant but unknown (i.e., a vector of parameters), we can condition everything in the basic and transformed models on d and proceed with estimation as described in Section 3 using an initial set of starting values for the elements of d. For a particular set of values for the elements of d, we can estimate the transformed model and hence the conditional expected efficiencies $\mu_\eta(Z_i, d) = \mathbb{E}(\eta_i \mid Z_i, d)$. Using iterative, numerical methods one can then optimize some criterion function $\mathcal{C}(\widehat{\mu}_\eta(Z_i, d) \mid i = 1, \ldots, n)$ with respect to d.

We caution that this approach would involve substantial computational burden since on each iteration, a new model would have to be estimated, requiring optimization of bandwidths on each iteration. Nonetheless, with increasingly widespread access to high performance computing systems in the form of clusters, grids or cloud computing systems allowing parallel computation, the approach is possible if one insists on a data-driven direction vector. One would also have to assume that the "true" direction vector is the one that optimizes the chosen criterion function.

E Monte Carlo Evidence

E.1 Experimental Framework

Here we describe results from a series of Monte Carlo experiments to gain insight on how well our estimation approach works in terms of recovering information about the true, underlying model. We consider various sample sizes n and various numbers of inputs (p) and outputs (q). We also consider various noise-to-signal ratios.

In order to simulate a frontier Ψ^∂ in settings with multiple inputs and multiple outputs, consider a sphere of radius one centered at the origin in $\mathbb{R}^{p+q}$. Then reflect the negative part of the q output axes and the positive part of the p input axes around the origin so that only $1/2^{p+q}$ of the sphere lying in $\mathbb{R}^p_- \times \mathbb{R}^q_+$ remains. Then shift the remaining surface along the p input axes in the positive direction by one unit so that the frontier Ψ^∂ lies in $\mathbb{R}^{p+q}_+$. In two dimensions with $p = q = 1$, this results in a frontier corresponding to the upper-left, northwest quarter of a circle centered at $(1, 0)$.

In each experiment, for given values of p and q, we first use the method of Muller (1959) and Marsaglia (1972) to draw $(p+q)$-tuples $a_i = (a_{p,i}, a_{q,i})$ uniformly distributed on the surface of a closed $(p+q)$-ball centered at the origin with radius one for $i = 1, \ldots, n$. The draws are independent, so that $a_i \perp\!\!\!\perp a_j$ for $i \neq j$. For each i we then set $X_i^\partial = (1 - |a_{p,i}|)$

and $Y_i^\partial = |a_{q,i}|$ to obtain efficient input-output pairs $(X_i^\partial, Y_i^\partial)$. For the direction vector we use $d = \widetilde{d}/||\widetilde{d}||$ with $\widetilde{d} = \begin{bmatrix} -i_p' & i_q' \end{bmatrix}'$ where i_p and i_q are column vectors of ones with lengths p and q, respectively. We normalize d to have length equal to one, but this is not required and has no bearing on the results of our experiments.

We consider two sets of experiments. Section E.2 discusses experments where inefficiency is distributed half-normal as in Assumption 2.5. Section E.3 discusses experments where inefficiency is distributed exponential, but the half-normal specification in Assumption 2.5 is maintained for estimation. The second set of experiments discussed in Secitn E.3 serves as a robustness check.

E.2 Half-Normal Inefficiency

We first consider homoskedastic deviations from the frontier. We set $\sigma_\eta = 0.5$ and draw $\eta_i \sim N^+(0, \sigma_\eta^2)$ for each $i = 1, \ldots, n$. For a given value of $\rho \geq 0$, we set

$$\sigma_\epsilon = \rho \left[(\pi - 2)/\pi\right]^{1/2} \sigma_\eta \tag{E.1}$$

and then draw $\epsilon_i \sim N(0, \sigma_\epsilon^2)$. The value of ρ serves to control the noise-to-signal ratio, i.e., the ratio of the standard deviations of the efficiency η_i and the noise ϵ_i.

Using the direction vector d to construct the rotation matrix R_d defined in (2.10), the efficient input-output pairs $(X_i^\partial, Y_i^\partial)$ are transformed to (z, u)-space by computing $(Z_i^\partial, U_i^\partial)$ via (2.20). We compute then construct "observed values" (Z_i, U_i) by setting $Z_i = Z_i^\partial$ and computing $U_i = U_i^\partial + \xi_i$ where $\xi_i = \epsilon_i - \eta_i$. Note that this is identical to computing "observed" values (X_i, Y_i) in (x, y)-space from the $(X_i^\partial, Y_i^\partial)$ using (2.4) and then computing the (Z_i, U_i) using (2.11).

In each set of experiments, we consider sample sizes $n \in \{100,\ 200,\ 400,\ 800,\ 1600,\ 3200,\ 6400\}$ and noise-to-signal ratios $\rho \in \{0.0,\ 0.25,\ 0.50,\ 0.75,\ 1.0,\ 2.0\}$. In addition, we consider settings with two to six dimensions with $(p, q) \in \{(1, 1),\ (1, 2),\ (2, 2),\ (2, 3),\ (3, 3)\}$. Each experiment with fixed values of n, ρ and (p, q) consists of 1,000 Monte Carlo trials.

For our first set of experiments, with homoskedastic noise and inefficiency, we apply both the SVKZ or HMS estimator described above to obtain two sets of estimates $\widehat{U}_i^\partial$ of U_i^∂ for $i = 1, \ldots, n$. For each Monte Carlo trial, we compute an estimate of integrated mean square error (IMSE) given by

$$\widehat{\text{IMSE}}_m = \sum_{i=1}^{n} \left(\widehat{U}_i^\partial - U_i^\partial \right)^2, \tag{E.2}$$

$m = 1 \ldots, 1000$ for either set of estimates. Table E.1 reports the sample means of these estimates (for both the SVKZ and HMS estimators) over the 1,000 Monte Carlo trials in each of the resulting $6 \times 7 \times 5 = 210$ experiments.

The results in Table E.1 indicate that for both estimators, IMSE decreases with sample size, and increases with both number of dimensions and the noise-to-signal ratio as expected. Moreover, when $\rho \leq 1$, the HMS estimator results in smaller IMSE than the SVKZ estimator in all but a few cases. To check whether the differences in IMSE between the two estimators are significant, IMSE for the SVKZ estimator for given value of n, p and q is subtracted from the corresponding IMSE for the HMS estimator for each Monte Carlo trial, and this difference is divided by the square root of the variance of the differences across Monte Carlo trials divided by the number of trials. By the Lindeberg-Levy central limit theorem, the resulting quantity is approximately distributed $N(0, 1)$, allowing two-sided tests of significance. The differences are shown in Table E.2, with asterisks indicating significance at 90, 95 or 99 percent levels. The differences in E.2 are significant at 90-percent or better in 179 of 210 cases. In addition, all of the significant differences are negative when $\rho \leq 1$. For $\rho = 2$, the differences are positive in many cases. In all cases, the differences in Table E.2 become smaller as sample size increases. When $\rho = 2$, the differences eventually become negative and significant when $n = 6400$, except in the case with 6 dimensions.

The next round of experiments allow for heteroskedasticity in both the inefficiency as well as the noise processes. We consider the same values of the noise-to-signal ratio ρ, dimensions (p, q) and sample size n as in the previous set of experiments, We again perform

1,000 Monte Carlo trials for each experiment, and the true frontier remains the same as before. We again focus on IMSE of the frontier estimates. In these experiments, we generate efficient observations $(X_i^\partial, Y_i^\partial)$ in (x,y)-space and transform these to efficient observations (Z_i, U_i^∂) exactly as in the first set of experiments. But instead of drawing $\eta_i \sim N^+(0, \sigma_\eta^2)$ as before, we draw $\eta_i \sim N^+\left(0,\ (0.5/(1+\exp(Z_i'\alpha)) + c_{pq})^2\right)$ where α is an $(r \times 1)$ vector of coefficients equal to 1 and c_{pq} is a constant depending on p and q as described below. Hence the mean and variance of η_i is made to depend on the elements of Z_i, and hence on all of the inputs and outputs in (x,y)-space. Similarly, for the noise we draw $\epsilon_i \sim N\left(0,\ (0.75 + (0.5/(1+Z_1\alpha)^2)\sigma_\epsilon^2\right)$.

To facilitate comparison of results from these experiments comparable with the results with no heteroskedasticity in Table E.1, for each pair (p,q) we first estimate the expectation of the factor $0.5/(1+\exp(Z_i\alpha))$ by simulating one billion draws of efficient input-output pairs $(X_i^\partial, Y_i^\partial)$, transforming these to (Z_i, U_i^∂) by multiplying by the rotation matrix R_d, and then computing the sample mean of the values $0.5/(1+\exp(Z_1\alpha))$, $i = 1, \ldots, 10^9$. We then subtract this mean from 0.5, and use the resulting value for σ_η. This yields values $c_{pq} = 0.1684$, 0.1288, 0.1056, 0.0747 and 0.0611 corresponding to $(p,q) = (1,1)$, (1,2), (2,2), (2,3) and (3,3), respectively, ensuring that $\mathbb{E}(\sigma_\eta(Z_i))$ is approximately 0.5 in each case. We then let σ_ϵ be determined by ρ as in (E.1). Finally, we use the draws of η_i and ϵ_i to construct "observed" data (X_i, Y_i) and (Z_i, U_i) as before.

Table E.3 shows our Monte Carlo estimates of IMSE for frontier estimates under heteroskedastic inefficiency and noise. Overall, we find the same patterns as before: for both estimators, IMSE decreases with increasing sample size, and increases with number of dimensions and with the noise-to-signal ratio ρ. Analogous to Table E.2, Table E.4 shows differences between average IMSE for the SVKZ and HMS estimators from Table E.3. The results indicate that the HMS estimator has significantly smaller IMSE than the SVKZ estimator in all but 15 of the 175 cases represented in Table E.3 where $\rho \leq 1$. Among these 15 cases, there are none where IMSE for the SVKZ estimator is significantly less than that for the HMS estimator. When $\rho = 2.0$, the IMSE estimator yields significantly smaller IMSE on average than the HMS estimator in 29 of 35 cases.

Comparing the results with heteroskedasticity in Table E.3 with the corresponding results in Table E.1 where there is no heteroskedasticity, we find mixed results. TableE.5 gives differences in average IMSE between Tables E.3 and E.1 for the SVKZ estimator, with positive signs indicating that IMSE is larger when there is no heteroskedasticity. Table E.6 gives similar differences for the HMS estimator. The results in both tables E.5 and E.6 indicate that some of the differences are statistically significant, but are numerically small.

Using the same experiments to produce the results in Tables E.1 and E.3, we also examine error in estimation of derivatives of expected inefficiency with respect to inputs and outputs in the original (x, y)-space. In our experiments, we consider marginal effects at a fixed point (x, y) on the frontier with the p elements of x determined by the median input level (which is the same for each of p dimensions) and the q elements of y equal to $[1-(x-1)'(x-1)]^{1/2}$. From (3.5) we have

$$\frac{\partial \mu_\eta(z)}{\partial z} = \sqrt{\frac{2}{\pi}} \frac{\partial \sigma_\eta(z)}{\partial z}. \tag{E.3}$$

For the SVKZ estimator of $\sigma_\eta(z)$ given in (3.8), we have

$$\frac{\partial \widehat{\sigma}_\eta(z)}{\partial z} = \begin{cases} \frac{1}{3}\left(\frac{\pi}{2}\right)^{1/2}\left(\frac{\pi}{\pi-4}\right)\left[\left(\frac{\pi}{2}\right)^{1/2}\left(\frac{pi}{\pi-4}\right)\widehat{r}_3(z)\right]^{-2/3}\frac{\partial \widehat{r}_3(z)}{\partial z} & \text{if } \widehat{r}_3(z) \le 0; \\ 0 & \text{otherwise.} \end{cases} \tag{E.4}$$

For the HMS estimator of $\sigma_\eta(z)$ given in (3.15), we have

$$\frac{\partial \widehat{\sigma}_\eta(z)}{\partial z} = \begin{cases} -\left(\frac{1}{3a_3^+}\right)\left[-\frac{1}{a_3^+}\widehat{r}_3(z)\right]^{-2/3}\frac{\partial \widehat{r}_3(z)}{\partial z} & \text{if } \widehat{r}_3(z) \le 0; \\ \left(\frac{1}{3a_3^-}\right)\left[\frac{1}{a_3^-}\widehat{r}_3(z)\right]^{-2/3}\frac{\partial \widehat{r}_3(z)}{\partial z} & \text{otherwise.} \end{cases} \tag{E.5}$$

Estimates $\widehat{r}_3(z)$ and $\partial\widehat{r}_3(z)/\partial z$ are obtained by LLLS estimation in the regression of cubed residuals $\widehat{\xi}_i$ on Z_i as discussed above in Section 3. Substituting these estimates into (E.3) and (E.5) give SVKZ and HMS estimates $\partial\widehat{\sigma}_\eta(z)/\partial z$, and substituting these into (A.10) give

SVKZ and HMS estimates of the partial of expected inefficiency with respect to inputs and outputs in the original (x, y)-space.

In each experiment, on each of 1,000 Monte Carlo trials, we compute $(p + q)$ estimates $\widehat{\mu}_\eta(w^\partial)/\partial w^\partial$, using either the SVKZ or HMS estimators, and then we compute the average squared error $\frac{1}{p+q}\left(\frac{\partial\widehat{\mu}_\eta(w^\partial)}{\partial w^\partial} - \frac{\partial\mu_\eta(w^\partial)}{\partial w^\partial}\right)'\left(\frac{\partial\widehat{\mu}_\eta(w^\partial)}{\partial w^\partial} - \frac{\partial\mu_\eta(w^\partial)}{\partial w^\partial}\right)$. Table E.7 reports medians of these average squared errors for each of our 210 Monte Carlo experiments.[2] The results in Table E.7 are similar in magnitude to those in Tables E.1–E.3. As expected, squared error increases with increasing dimensionality, and diminishes with increasing sample size, although this effect is small when $p = q = 1$.

To examine marginal products, we consider the same fixed point as in the previous exercise. Recalling the discussion at the beginning of Section E, note that a point (x, y) on the frontier must satisfy $(x - 1)'(x - 1) + y'y = 1$. For a fixed point (x, y) and $j \neq k$, there are pq distinct marginal products $\frac{\partial y^k}{\partial x^j} = (1 - x^j)/y^k$, where x^j and y^k denote the jth and kth elements of x and j, respectively. To obtain estimates of marginal products, recall that estimates of $\partial\phi(z)/\partial z$ can be obtained as discussed in Section 3.1, and these can be used to estimate the Jacobian matrix in (A.2). Substituting estimates of the Jacobian into (A.6) and solving for the r-vector c allows marginal products to be estimated using (A.7).

On each Monte Carlo trial, we compute pq marginal products using either the SVKZ or HMS estimators, subtract the corresponding true values, and then compute the sum of squares of the resulting differences divided by $(p \times q)$. Table E.8 reports median average squared errors across 1,000 Monte Carlo trials for each of our 210 experiments, analogous to Table E.7. The medians in Table E.8 are necessarily larger than the corresponding values in Table E.7 due to the fact that $\widehat{r}_1(z)$ is added to $\widehat{\mu}_z(z)$ to obtain the frontier estimate $\widehat{\phi}(z)$

[2]Estimation of derivatives is necessarily more difficult than estimation of conditional means. Among 1,000 Monte Carlo trials, we sometimes find in a hand-full of cases derivative estimates that are implausibly large; in other words, the distribution of our squared errors across 1,000 trials is skewed to the right. Consequently, we report medians in Table E.7 to provide a more robust, stable measure than would be provided by the mean of the squared errors.

in (3.11). Nonetheless, the squared errors in Table E.8 are numerically small provided the number of dimensions is not too large.

E.3 Exponential Inefficiency

In order to check the robustness of our simulation results, we repeat both rounds of experiments described above in in Section E.2, replacing the half-normal distribution for inefficiency with an exponential distribution when simulating data, but maintaining Assumption 2.5 for purposes fo estimation. One should expect to pay a price in terms of estimation error for mis-specifying the inefficiency process, but due to our local estimation, the results indicate the price is small as shown below.

In the first group of experiments, both inefficiency and noise are homoskedastic as in Section E.2. For each $i = 1, \ldots, n$ we draw η_i from the exponential density $f(\eta \mid \omega) = \omega \exp(-\omega\eta)$ instead of the half-normal distribution $N^+(0, \sigma_\eta^2)$ used in Section E.2. In order to facilitate comparison with the results in Table E.1, we set $\omega = \sqrt{\pi}(\sqrt{2}\sigma_\eta)^{-1}$, with $\sigma_\eta = 0.5$ as before, so that the exponential inefficiencies have the same mean as the half-normal inefficiencies in Section E.2. For the same values of ρ as in Section E.2, we then set $\sigma_\epsilon = \rho\sigma_\eta\sqrt{2}/\sqrt{\pi}$ and draw ϵ_i from $N(0, \sigma_\epsilon^2)$ for each $i = 1, \ldots, n$. As in Section E.2, ρ controls the noise-to-signal ratio in terms of the ratios of standard deviations of the noise and the inefficiency processes.

In the second group of experiments, both inefficiency and noise are heteroskedastic, exponentially-distributed. For each $i = 1, \ldots, n$, we compute $\sigma_\eta(Z_i) = (0.5/(1+\exp(Z_i\alpha))+ c_{p,q}$ and $\sigma_\epsilon(Z_i) = \rho\sigma_\eta(Z_i)\sqrt{2}/\sqrt{\pi}$ with ρ controlling the noise-to-signal ratio as before. The constants $c_{p,q}$ are described in Section E.2 of the paper, and ensure that the expectation of $\sigma_\eta(Z_i)$ equals 0.5 to facilitate comparison with the results in Tables E.1 and E.9. We then set $\omega(Z_i) = (\sqrt{2}/\sqrt{\pi})\sigma_\eta(Z_i)$ and draw η_i from the exponential density with parameter $\omega(Z_i)$, and draw ϵ_i from $N(0, \sigma_\epsilon(Z_i))$. The draws η_i and ϵ_i are used to construct (X_i, Y_i) and

(Z_i, U_i) as described in Section E.2.

Results from the two sets of experiments with exponentially-distributed inefficiency are shown in Tables E.9–E.12, analogous to the results from experiments with half-normal inefficiency shown in Tables E.1–E.8 in the main paper. Of course, one should expect to pay a price for mis-specification of the inefficiency process. To facilitate comparison of the experiments where inefficiency is exponential with the corresponding experiment where inefficiency is half-normal, Tables E.13–E.16 show differences in IMSEs and median average square errors between Tables E.9–E.12 (where inefficiency is distributed exponential) and Tables E.1–E.8 (where inefficiency is distributed half-normal). Negative signs in Tables E.13–E.16 indicate smaller errors with the mis-specified inefficiency, but occur only in a few cases in each table. However, while most of the differences in Tables E.13–E.16 are positive, most are also of small magnitude. For example, the median and mean values of the differences (between Tables E.9 and E.1) shown in Table E.13 are 0.0447 and 0.0593, respectively. Similarly, the median and mean values of the differences (between Tables E.10 and E.3) shown in Table E.14 are 0.0474 and 0.0629, respectively. For purposes of estimating derivatives of expected efficiency, the differences between Tables E.11 and E.7 shown in Table E.15 are even smaller, with median and mean values of 0.0247 and 0.0411, respectively. For purposes of estimating marginal products, mis-specification of the inefficiency process is less costly still—the differences between Tables E.12 and E.8 shown in Table E.16 have median and mean values of 0.0047 and −0.0325 (respectively), and the differences are *negative* in 186 of 420 cases. While there is a price to pay for mis-specifying exponential inefficiency as half-normal, the results suggest that in most cases the price is not high, especially for purposes of estimating marginal products and derivatives of expected inefficiency.

Table E.1: IMSE for Frontier Estimates with No Heteroskedasticity

		SVKZ					HMS				
ρ	n	(1,1)	(1,2)	(2,2)	(2,3)	(3,3)	(1,1)	(1,2)	(2,2)	(2,3)	(3,3)
0.00	100	0.0193	0.0369	0.0461	0.0548	0.0632	0.0161	0.0274	0.0344	0.0388	0.0437
	200	0.0094	0.0227	0.0288	0.0337	0.0401	0.0082	0.0187	0.0228	0.0263	0.0293
	400	0.0043	0.0125	0.0176	0.0205	0.0239	0.0042	0.0113	0.0153	0.0170	0.0192
	800	0.0021	0.0067	0.0115	0.0131	0.0140	0.0021	0.0066	0.0111	0.0121	0.0125
	1600	0.0011	0.0036	0.0077	0.0094	0.0098	0.0010	0.0036	0.0077	0.0092	0.0094
	3200	0.0006	0.0020	0.0052	0.0073	0.0076	0.0006	0.0020	0.0051	0.0073	0.0075
	6400	0.0003	0.0012	0.0033	0.0059	0.0065	0.0003	0.0012	0.0033	0.0059	0.0065
0.25	100	0.0221	0.0407	0.0497	0.0578	0.0668	0.0174	0.0291	0.0366	0.0409	0.0462
	200	0.0104	0.0251	0.0322	0.0366	0.0427	0.0090	0.0199	0.0245	0.0279	0.0306
	400	0.0050	0.0142	0.0194	0.0224	0.0263	0.0047	0.0125	0.0164	0.0180	0.0205
	800	0.0023	0.0073	0.0126	0.0143	0.0149	0.0023	0.0072	0.0119	0.0129	0.0131
	1600	0.0012	0.0040	0.0084	0.0101	0.0104	0.0012	0.0039	0.0084	0.0099	0.0099
	3200	0.0006	0.0022	0.0056	0.0078	0.0081	0.0006	0.0022	0.0056	0.0078	0.0080
	6400	0.0003	0.0013	0.0036	0.0063	0.0070	0.0003	0.0013	0.0036	0.0063	0.0070
0.50	100	0.0323	0.0516	0.0601	0.0690	0.0776	0.0216	0.0346	0.0438	0.0508	0.0558
	200	0.0150	0.0340	0.0419	0.0459	0.0515	0.0118	0.0240	0.0293	0.0329	0.0366
	400	0.0071	0.0204	0.0261	0.0293	0.0336	0.0063	0.0159	0.0198	0.0214	0.0250
	800	0.0031	0.0100	0.0171	0.0185	0.0195	0.0031	0.0094	0.0147	0.0155	0.0159
	1600	0.0015	0.0053	0.0110	0.0130	0.0132	0.0015	0.0052	0.0106	0.0121	0.0119
	3200	0.0008	0.0029	0.0072	0.0096	0.0099	0.0008	0.0029	0.0072	0.0095	0.0097
	6400	0.0004	0.0016	0.0047	0.0078	0.0083	0.0004	0.0016	0.0047	0.0078	0.0083
0.75	100	0.0490	0.0672	0.0755	0.0838	0.0903	0.0315	0.0503	0.0636	0.0746	0.0808
	200	0.0257	0.0507	0.0574	0.0606	0.0663	0.0178	0.0328	0.0409	0.0471	0.0536
	400	0.0128	0.0337	0.0396	0.0431	0.0466	0.0097	0.0216	0.0265	0.0297	0.0339
	800	0.0053	0.0180	0.0281	0.0293	0.0300	0.0049	0.0137	0.0198	0.0210	0.0217
	1600	0.0024	0.0088	0.0181	0.0204	0.0207	0.0024	0.0081	0.0153	0.0164	0.0161
	3200	0.0012	0.0045	0.0112	0.0145	0.0148	0.0012	0.0045	0.0108	0.0134	0.0132
	6400	0.0006	0.0025	0.0071	0.0112	0.0116	0.0006	0.0025	0.0071	0.0111	0.0113
1.00	100	0.0675	0.0828	0.0896	0.0986	0.1055	0.0571	0.0856	0.1057	0.1212	0.1320
	200	0.0456	0.0691	0.0741	0.0770	0.0819	0.0303	0.0548	0.0686	0.0802	0.0883
	400	0.0270	0.0528	0.0567	0.0602	0.0636	0.0166	0.0337	0.0408	0.0488	0.0564
	800	0.0117	0.0357	0.0468	0.0472	0.0463	0.0088	0.0214	0.0286	0.0320	0.0342
	1600	0.0047	0.0192	0.0351	0.0357	0.0352	0.0042	0.0141	0.0222	0.0232	0.0238
	3200	0.0022	0.0090	0.0227	0.0270	0.0272	0.0022	0.0081	0.0178	0.0196	0.0192
	6400	0.0010	0.0043	0.0129	0.0200	0.0202	0.0010	0.0043	0.0123	0.0174	0.0168
2.00	100	0.1262	0.1416	0.1542	0.1691	0.1783	0.4107	0.4885	0.5573	0.5886	0.6284
	200	0.1079	0.1217	0.1317	0.1368	0.1435	0.2902	0.3862	0.4266	0.4731	0.4924
	400	0.0941	0.1087	0.1107	0.1164	0.1207	0.1980	0.2651	0.2907	0.3482	0.3780
	800	0.0811	0.1031	0.1045	0.1060	0.1052	0.1195	0.2014	0.2294	0.2542	0.2745
	1600	0.0669	0.1030	0.1041	0.0976	0.0941	0.0662	0.1474	0.1722	0.1728	0.1898
	3200	0.0431	0.0978	0.1085	0.0990	0.0944	0.0300	0.0959	0.1302	0.1298	0.1387
	6400	0.0227	0.0820	0.1168	0.1093	0.0967	0.0134	0.0538	0.1018	0.1051	0.0985

Table E.2: Differences in IMSE, HMS versus SVKZ, No Heteroskedasticity

ρ	n	(1,1)	(1,2)	(2,2)	(2,3)	(3,3)
0.00	100	−0.003244***	−0.009575***	−0.011749***	−0.016052***	−0.019432***
	200	−0.001154***	−0.004065***	−0.005942***	−0.007412***	−0.010825***
	400	−0.000105***	−0.001175***	−0.002346***	−0.003520***	−0.004755***
	800	0.000000	−0.000128***	−0.000441***	−0.001049***	−0.001549***
	1600	−0.000002	−0.000001	−0.000055***	−0.000177***	−0.000423***
	3200	0.000000	0.000000	−0.000004	−0.000018***	−0.000060***
	6400	0.000000	0.000000	−0.000001	−0.000001*	−0.000012**
0.25	100	−0.004672***	−0.011608***	−0.013121***	−0.016972***	−0.020602***
	200	−0.001361***	−0.005174***	−0.007784***	−0.008741***	−0.012084***
	400	−0.000270***	−0.001723***	−0.002976***	−0.004444***	−0.005807***
	800	−0.000014	−0.000158***	−0.000699***	−0.001375***	−0.001788***
	1600	−0.000006	−0.000013*	−0.000094***	−0.000253***	−0.000493***
	3200	0.000000	0.000000	−0.000002	−0.000015***	−0.000076***
	6400	0.000000	0.000000	−0.000002	−0.000001**	−0.000018**
0.50	100	−0.010666***	−0.016970***	−0.016289***	−0.018240***	−0.021777***
	200	−0.003191***	−0.009998***	−0.012592***	−0.013073***	−0.014949***
	400	−0.000742***	−0.004500***	−0.006288***	−0.007929***	−0.008649***
	800	−0.000057**	−0.000672***	−0.002367***	−0.002985***	−0.003615***
	1600	−0.000009	−0.000079***	−0.000413***	−0.000833***	−0.001323***
	3200	0.000000	−0.000001	−0.000020***	−0.000088***	−0.000241***
	6400	0.000000	0.000000	0.000000	−0.000017*	−0.000042***
0.75	100	−0.017536***	−0.016928***	−0.011953***	−0.009232***	−0.009488***
	200	−0.007948***	−0.017874***	−0.016524***	−0.013497***	−0.012658***
	400	−0.003082***	−0.012089***	−0.013062***	−0.013429***	−0.012683***
	800	−0.000345***	−0.004297***	−0.008249***	−0.008294***	−0.008326***
	1600	−0.000043***	−0.000674***	−0.002811***	−0.004056***	−0.004561***
	3200	0.000000	−0.000015**	−0.000366***	−0.001110***	−0.001606***
	6400	0.000000	0.000000	−0.000007***	−0.000177***	−0.000359***
1.00	100	−0.010354***	0.002823	0.016089***	0.022600***	0.026479***
	200	−0.015329***	−0.014297***	−0.005488***	0.003168**	0.006327***
	400	−0.010336***	−0.019091***	−0.015948***	−0.011386***	−0.007143***
	800	−0.002917***	−0.014280***	−0.018196***	−0.015210***	−0.012096***
	1600	−0.000456***	−0.005124***	−0.012921***	−0.012535***	−0.011378***
	3200	−0.000003	−0.000839***	−0.004972***	−0.007344***	−0.008035***
	6400	0.000000	−0.000027***	−0.000634***	−0.002547***	−0.003389***
2.00	100	0.284463***	0.346932***	0.403054***	0.419524***	0.450096***
	200	0.182323***	0.264428***	0.294969***	0.336303***	0.348912***
	400	0.103878***	0.156422***	0.179949***	0.231749***	0.257277***
	800	0.038412***	0.098306***	0.124905***	0.148143***	0.169266***
	1600	−0.000738	0.044403***	0.068119***	0.075144***	0.095688***
	3200	−0.013134***	−0.001912	0.021663***	0.030841***	0.044368***
	6400	−0.009266***	−0.028138***	−0.015012***	−0.004225***	0.001889

NOTE: One, two or three asterisks indicate siginifcant differences from zero at 90, 95 or 99-percent levels.

Table E.3: IMSE for Frontier Estimates with Heteroskedastic Inefficiency and Noise

		SVKZ					HMS				
ρ	n	(1,1)	(1,2)	(2,2)	(2,3)	(3,3)	(1,1)	(1,2)	(2,2)	(2,3)	(3,3)
0.00	100	0.0207	0.0356	0.0461	0.0546	0.0655	0.0179	0.0298	0.0349	0.0413	0.0444
	200	0.0099	0.0229	0.0292	0.0332	0.0400	0.0089	0.0196	0.0244	0.0263	0.0302
	400	0.0050	0.0135	0.0185	0.0207	0.0236	0.0045	0.0119	0.0162	0.0175	0.0192
	800	0.0023	0.0071	0.0120	0.0134	0.0144	0.0022	0.0067	0.0111	0.0123	0.0128
	1600	0.0012	0.0038	0.0079	0.0093	0.0097	0.0012	0.0037	0.0077	0.0090	0.0092
	3200	0.0006	0.0021	0.0051	0.0071	0.0074	0.0006	0.0021	0.0051	0.0071	0.0073
	6400	0.0003	0.0012	0.0033	0.0057	0.0063	0.0003	0.0012	0.0033	0.0057	0.0063
0.25	100	0.0230	0.0391	0.0488	0.0572	0.0693	0.0193	0.0312	0.0366	0.0429	0.0470
	200	0.0111	0.0255	0.0316	0.0364	0.0434	0.0100	0.0211	0.0254	0.0279	0.0323
	400	0.0055	0.0147	0.0199	0.0229	0.0257	0.0050	0.0130	0.0171	0.0189	0.0206
	800	0.0025	0.0080	0.0131	0.0147	0.0154	0.0024	0.0074	0.0120	0.0132	0.0135
	1600	0.0013	0.0042	0.0086	0.0100	0.0103	0.0013	0.0040	0.0083	0.0096	0.0097
	3200	0.0006	0.0023	0.0055	0.0076	0.0079	0.0006	0.0023	0.0055	0.0075	0.0078
	6400	0.0003	0.0013	0.0036	0.0061	0.0067	0.0003	0.0013	0.0036	0.0061	0.0067
0.50	100	0.0321	0.0503	0.0591	0.0670	0.0789	0.0238	0.0370	0.0444	0.0522	0.0582
	200	0.0157	0.0334	0.0405	0.0458	0.0537	0.0129	0.0258	0.0301	0.0337	0.0398
	400	0.0078	0.0204	0.0266	0.0302	0.0331	0.0067	0.0165	0.0209	0.0234	0.0256
	800	0.0035	0.0108	0.0171	0.0195	0.0203	0.0033	0.0096	0.0147	0.0161	0.0168
	1600	0.0017	0.0057	0.0112	0.0128	0.0132	0.0017	0.0054	0.0105	0.0117	0.0119
	3200	0.0008	0.0030	0.0071	0.0094	0.0098	0.0008	0.0030	0.0070	0.0092	0.0095
	6400	0.0004	0.0017	0.0046	0.0075	0.0081	0.0004	0.0017	0.0046	0.0075	0.0080
0.75	100	0.0486	0.0657	0.0749	0.0810	0.0930	0.0335	0.0530	0.0646	0.0773	0.0858
	200	0.0268	0.0485	0.0557	0.0609	0.0681	0.0194	0.0355	0.0421	0.0481	0.0565
	400	0.0135	0.0331	0.0391	0.0439	0.0475	0.0107	0.0233	0.0285	0.0335	0.0382
	800	0.0060	0.0184	0.0269	0.0305	0.0304	0.0052	0.0146	0.0203	0.0218	0.0230
	1600	0.0028	0.0095	0.0179	0.0200	0.0206	0.0027	0.0084	0.0151	0.0160	0.0162
	3200	0.0013	0.0049	0.0113	0.0141	0.0147	0.0013	0.0046	0.0105	0.0129	0.0131
	6400	0.0006	0.0026	0.0071	0.0109	0.0115	0.0006	0.0025	0.0070	0.0106	0.0110
1.00	100	0.0668	0.0818	0.0902	0.0967	0.1073	0.0580	0.0878	0.1043	0.1234	0.1355
	200	0.0450	0.0671	0.0733	0.0772	0.0840	0.0341	0.0586	0.0707	0.0799	0.0932
	400	0.0272	0.0518	0.0566	0.0610	0.0637	0.0182	0.0359	0.0444	0.0546	0.0613
	800	0.0127	0.0347	0.0453	0.0474	0.0465	0.0097	0.0230	0.0297	0.0332	0.0360
	1600	0.0057	0.0200	0.0342	0.0353	0.0346	0.0049	0.0147	0.0225	0.0230	0.0234
	3200	0.0023	0.0095	0.0223	0.0263	0.0260	0.0022	0.0083	0.0172	0.0191	0.0188
	6400	0.0011	0.0046	0.0132	0.0196	0.0196	0.0011	0.0044	0.0120	0.0167	0.0161
2.00	100	0.1269	0.1424	0.1563	0.1669	0.1822	0.4104	0.4964	0.5452	0.5997	0.6247
	200	0.1066	0.1252	0.1302	0.1358	0.1467	0.2975	0.3933	0.4452	0.4688	0.5242
	400	0.0930	0.1094	0.1120	0.1178	0.1223	0.1958	0.2762	0.3186	0.3509	0.3954
	800	0.0824	0.1032	0.1057	0.1055	0.1055	0.1277	0.2026	0.2331	0.2620	0.2855
	1600	0.0677	0.1019	0.1042	0.0968	0.0948	0.0724	0.1493	0.1713	0.1817	0.1895
	3200	0.0457	0.0955	0.1072	0.0991	0.0934	0.0346	0.0951	0.1324	0.1322	0.1392
	6400	0.0256	0.0791	0.1126	0.1073	0.0934	0.0161	0.0545	0.1013	0.1048	0.0976

Table E.4: Differences in IMSE, HMS versus SVKZ, Heteroskedastic Inefficiency and Noise

ρ	n	(1,1)	(1,2)	(2,2)	(2,3)	(3,3)
0.00	100	−0.002813***	−0.005766***	−0.011150***	−0.013373***	−0.021110***
	200	−0.000963***	−0.003373***	−0.004775***	−0.006889***	−0.009817***
	400	−0.000427***	−0.001582***	−0.002285***	−0.003193***	−0.004339***
	800	−0.000056***	−0.000377***	−0.000871***	−0.001132***	−0.001603***
	1600	0.000000	−0.000104***	−0.000254***	−0.000306***	−0.000496***
	3200	−0.000003	−0.000010***	−0.000040***	−0.000039***	−0.000131***
	6400	0.000000	0.000000	−0.000003***	−0.000010**	−0.000022***
0.25	100	−0.003743***	−0.007947***	−0.012171***	−0.014287***	−0.022296***
	200	−0.001113***	−0.004460***	−0.006191***	−0.008480***	−0.011055***
	400	−0.000540***	−0.001795***	−0.002853***	−0.004026***	−0.005116***
	800	−0.000070***	−0.000567***	−0.001114***	−0.001541***	−0.001904***
	1600	−0.000001	−0.000132***	−0.000325***	−0.000430***	−0.000581***
	3200	−0.000003	−0.000015***	−0.000058***	−0.000065***	−0.000176***
	6400	0.000000	0.000000	−0.000005***	−0.000008***	−0.000037***
0.50	100	−0.008305***	−0.013306***	−0.014689***	−0.014725***	−0.020692***
	200	−0.002850***	−0.007571***	−0.010445***	−0.012047***	−0.013845***
	400	−0.001091***	−0.003873***	−0.005626***	−0.006824***	−0.007486***
	800	−0.000232***	−0.001198***	−0.002376***	−0.003439***	−0.003492***
	1600	−0.000011***	−0.000321***	−0.000728***	−0.001019***	−0.001370***
	3200	−0.000006	−0.000048***	−0.000144***	−0.000209***	−0.000380***
	6400	0.000000	−0.000006***	−0.000019***	−0.000029***	−0.000076***
0.75	100	−0.015114***	−0.012690***	−0.010371***	−0.003768**	−0.007164***
	200	−0.007438***	−0.012982***	−0.013639***	−0.012773***	−0.011660***
	400	−0.002773***	−0.009827***	−0.010590***	−0.010387***	−0.009277***
	800	−0.000780***	−0.003817***	−0.006660***	−0.008681***	−0.007365***
	1600	−0.000160***	−0.001097***	−0.002841***	−0.003981***	−0.004421***
	3200	−0.000001	−0.000215***	−0.000736***	−0.001188***	−0.001617***
	6400	0.000000	−0.000033***	−0.000127***	−0.000265***	−0.000532***
1.00	100	−0.008824***	0.006066***	0.014108***	0.026786***	0.028212***
	200	−0.010938***	−0.008559***	−0.002612*	0.002615*	0.009202***
	400	−0.008980***	−0.015850***	−0.012159***	−0.006390***	−0.002354**
	800	−0.002965***	−0.011678***	−0.015565***	−0.014273***	−0.010475***
	1600	−0.000794***	−0.005358***	−0.011709***	−0.012254***	−0.011125***
	3200	−0.000085***	−0.001182***	−0.005098***	−0.007147***	−0.007214***
	6400	−0.000002	−0.000199***	−0.001147***	−0.002904***	−0.003488***
2.00	100	0.283525***	0.353977***	0.388882***	0.432786***	0.442439***
	200	0.190896***	0.268136***	0.315044***	0.332992***	0.377434***
	400	0.102743***	0.166809***	0.206615***	0.233049***	0.273093***
	800	0.045325***	0.099329***	0.127388***	0.156520***	0.179984***
	1600	0.004660**	0.047427***	0.067080***	0.084874***	0.094716***
	3200	−0.011095***	−0.000350	0.025187***	0.033054***	0.045881***
	6400	−0.009582***	−0.024610***	−0.011318***	−0.002496*	0.004231***

NOTE: One, two or three asterisks indicate siginifcant differences from zero at 90, 95 or 99-percent levels.

Table E.5: Differences in IMSE for SVKZ Estimator Due to Heteroskedasticity

ρ	n	(1,1)	(1,2)	(2,2)	(2,3)	(3,3)
0.00	100	0.001336	−0.001376	−0.000007	−0.000186	0.002361
	200	0.000544	0.000202	0.000420	−0.000503	−0.000127
	400	0.000625***	0.000947**	0.000850*	0.000214	−0.000348
	800	0.000164**	0.000394**	0.000516*	0.000251	0.000329
	1600	0.000120***	0.000152*	0.000160	−0.000049	−0.000074
	3200	0.000022	0.000073*	−0.000034	−0.000241***	−0.000124
	6400	−0.000001	0.000007	−0.000051	−0.000196***	−0.000253***
0.25	100	0.000940	−0.001569	−0.000981	−0.000589	0.002469
	200	0.000698	0.000436	−0.000677	−0.000225	0.000735
	400	0.000541**	0.000518	0.000549	0.000503	−0.000658
	800	0.000181**	0.000662***	0.000435	0.000445	0.000535
	1600	0.000140***	0.000224**	0.000150	−0.000075	−0.000115
	3200	0.000029	0.000085*	−0.000049	−0.000205**	−0.000119
	6400	0.000012	0.000020	−0.000058	−0.000235***	−0.000256***
0.50	100	−0.000149	−0.001250	−0.001005	−0.002062	0.001309
	200	0.000746	−0.000677	−0.001382	−0.000145	0.002146*
	400	0.000691*	−0.000006	0.000475	0.000898	−0.000503
	800	0.000391***	0.000749**	0.000014	0.001049*	0.000802
	1600	0.000208***	0.000431***	0.000258	−0.000202	−0.000005
	3200	0.000042	0.000137**	−0.000073	−0.000194	−0.000077
	6400	0.000029**	0.000039	−0.000051	−0.000281***	−0.000249**
0.75	100	−0.000476	−0.001518	−0.000583	−0.002797	0.002693
	200	0.001063	−0.002162	−0.001694	0.000261	0.001845
	400	0.000691	−0.000543	−0.000468	0.000751	0.000893
	800	0.000730***	0.000345	−0.001124	0.001201	0.000340
	1600	0.000437***	0.000734***	−0.000169	−0.000465	−0.000124
	3200	0.000040	0.000338***	0.000087	−0.000422	−0.000161
	6400	0.000066***	0.000087*	0.000053	−0.000366**	−0.000133
1.00	100	−0.000659	−0.000998	0.000585	−0.001918	0.001781
	200	−0.000604	−0.001933	−0.000804	0.000239	0.002112
	400	0.000200	−0.001044	−0.000160	0.000787	0.000110
	800	0.000946	−0.001029	−0.001476	0.000245	0.000229
	1600	0.000985***	0.000849	−0.000831	−0.000470	−0.000609
	3200	0.000158*	0.000496*	−0.000472	−0.000679	−0.001230*
	6400	0.000152***	0.000246**	0.000262	−0.000411	−0.000597
2.00	100	0.000667	0.000816	0.002061	−0.002140	0.003885
	200	−0.001299	0.003435	−0.001499	−0.000988	0.003278*
	400	−0.001039	0.000706	0.001246	0.001413	0.001534
	800	0.001258	0.000093	0.001168	−0.000585	0.000283
	1600	0.000791	−0.001125	0.000118	−0.000823	0.000745
	3200	0.002583	−0.002313	−0.001287	0.000127	−0.001018
	6400	0.002995**	−0.002856	−0.004210***	−0.002024	−0.003250**

NOTE: One, two or three asterisks indicate siginifcant differences from zero at 90, 95 or 99-percent levels.

Table E.6: Differences in IMSE for HMS Estimator Due to Heteroskedasticity

ρ	n	(1,1)	(1,2)	(2,2)	(2,3)	(3,3)
0.00	100	0.001767***	0.002434***	0.000592	0.002493**	0.000683
	200	0.000737**	0.000895	0.001588***	0.000019	0.000881
	400	0.000302*	0.000539*	0.000911**	0.000541	0.000068
	800	0.000108	0.000145	0.000086	0.000167	0.000275
	1600	0.000122***	0.000049	−0.000038	−0.000179	−0.000147
	3200	0.000019	0.000064	−0.000070	−0.000263***	−0.000196*
	6400	−0.000001	0.000007	−0.000053	−0.000206***	−0.000262***
0.25	100	0.001868***	0.002093**	−0.000032	0.002096**	0.000775
	200	0.000945***	0.001149**	0.000916	0.000036	0.001765**
	400	0.000271	0.000448	0.000672*	0.000921**	0.000033
	800	0.000125	0.000252	0.000019	0.000279	0.000419
	1600	0.000145***	0.000105	−0.000082	−0.000251	−0.000203
	3200	0.000026	0.000070	−0.000105	−0.000254**	−0.000219
	6400	0.000012	0.000019	−0.000062	−0.000242***	−0.000275***
0.50	100	0.002213**	0.002414**	0.000595	0.001452	0.002393*
	200	0.001087**	0.001750***	0.000765	0.000881	0.003250***
	400	0.000341	0.000621	0.001137**	0.002002***	0.000660
	800	0.000216*	0.000224	0.000005	0.000595*	0.000924**
	1600	0.000206***	0.000189	−0.000056	−0.000389*	−0.000052
	3200	0.000035	0.000090	−0.000197*	−0.000314**	−0.000216
	6400	0.000029**	0.000033	−0.000070	−0.000292***	−0.000283***
0.75	100	0.001946	0.002720*	0.001000	0.002667	0.005017**
	200	0.001573**	0.002731***	0.001191	0.000985	0.002843*
	400	0.001000**	0.001719***	0.002004***	0.003795***	0.004300***
	800	0.000295	0.000824**	0.000465	0.000815	0.001301**
	1600	0.000321***	0.000313	−0.000198	−0.000390	0.000016
	3200	0.000040	0.000138	−0.000283	−0.000499**	−0.000172
	6400	0.000066***	0.000054	−0.000067	−0.000455***	−0.000307*
1.00	100	0.000872	0.002244	−0.001397	0.002268	0.003515
	200	0.003787***	0.003806*	0.002072	−0.000315	0.004987**
	400	0.001554**	0.002197**	0.003630***	0.005784***	0.004899**
	800	0.000897**	0.001573***	0.001154	0.001183	0.001849*
	1600	0.000647***	0.000615*	0.000381	−0.000189	−0.000356
	3200	0.000077	0.000153	−0.000598*	−0.000482	−0.000408
	6400	0.000150***	0.000074	−0.000251	−0.000768***	−0.000695***
2.00	100	−0.000271	0.007861	−0.012111	0.011122	−0.003772
	200	0.007273	0.007144	0.018575	−0.004299	0.031801**
	400	−0.002174	0.011093	0.027912***	0.002714	0.017351*
	800	0.008172	0.001116	0.003650	0.007793	0.011001
	1600	0.006188*	0.001899	−0.000921	0.008907*	−0.000226
	3200	0.004622***	−0.000751	0.002237	0.002340	0.000494
	6400	0.002680***	0.000672	−0.000516	−0.000295	−0.000907

NOTE: One, two or three asterisks indicate siginifcant differences from zero at 90, 95 or 99-percent levels.

Table E.7: Median Average Square Error for Derivatives of Expected Efficiency $\mu_\eta(Z)$, with Heteroskedasticity

		SVKZ					HMS				
ρ	n	(1,1)	(1,2)	(2,2)	(2,3)	(3,3)	(1,1)	(1,2)	(2,2)	(2,3)	(3,3)
0.00	100	0.0182	0.0330	0.0458	0.0571	0.0663	0.0183	0.0350	0.0509	0.0642	0.0784
	200	0.0170	0.0284	0.0322	0.0364	0.0403	0.0170	0.0288	0.0325	0.0365	0.0418
	400	0.0163	0.0219	0.0223	0.0229	0.0256	0.0163	0.0219	0.0224	0.0230	0.0259
	800	0.0176	0.0185	0.0184	0.0168	0.0167	0.0176	0.0185	0.0184	0.0168	0.0167
	1600	0.0177	0.0176	0.0155	0.0124	0.0121	0.0177	0.0176	0.0155	0.0124	0.0121
	3200	0.0179	0.0164	0.0134	0.0100	0.0088	0.0179	0.0164	0.0134	0.0100	0.0088
	6400	0.0175	0.0164	0.0129	0.0087	0.0077	0.0175	0.0164	0.0129	0.0087	0.0077
0.25	100	0.0187	0.0393	0.0502	0.0624	0.0710	0.0197	0.0413	0.0565	0.0726	0.0954
	200	0.0179	0.0304	0.0355	0.0394	0.0468	0.0180	0.0307	0.0363	0.0405	0.0492
	400	0.0168	0.0229	0.0246	0.0256	0.0292	0.0168	0.0229	0.0246	0.0259	0.0295
	800	0.0176	0.0198	0.0191	0.0182	0.0181	0.0176	0.0198	0.0191	0.0183	0.0181
	1600	0.0172	0.0173	0.0158	0.0132	0.0130	0.0172	0.0173	0.0158	0.0132	0.0131
	3200	0.0177	0.0164	0.0137	0.0105	0.0092	0.0177	0.0164	0.0137	0.0105	0.0092
	6400	0.0177	0.0164	0.0130	0.0089	0.0081	0.0177	0.0164	0.0130	0.0089	0.0081
0.50	100	0.0219	0.0512	0.0659	0.0850	0.0843	0.0265	0.0635	0.0851	0.1130	0.1474
	200	0.0196	0.0400	0.0463	0.0545	0.0662	0.0200	0.0414	0.0494	0.0600	0.0764
	400	0.0178	0.0270	0.0332	0.0348	0.0399	0.0178	0.0273	0.0337	0.0360	0.0403
	800	0.0177	0.0218	0.0228	0.0242	0.0241	0.0177	0.0218	0.0228	0.0242	0.0243
	1600	0.0173	0.0181	0.0177	0.0159	0.0160	0.0173	0.0181	0.0177	0.0159	0.0160
	3200	0.0179	0.0170	0.0145	0.0124	0.0114	0.0179	0.0170	0.0145	0.0124	0.0114
	6400	0.0176	0.0163	0.0135	0.0098	0.0095	0.0176	0.0163	0.0135	0.0098	0.0095
0.75	100	0.0269	0.0573	0.0828	0.1034	0.1089	0.0462	0.1046	0.1483	0.2019	0.2829
	200	0.0264	0.0589	0.0644	0.0826	0.0928	0.0284	0.0731	0.0901	0.1056	0.1351
	400	0.0190	0.0362	0.0510	0.0558	0.0631	0.0193	0.0391	0.0558	0.0633	0.0726
	800	0.0180	0.0272	0.0324	0.0391	0.0403	0.0180	0.0273	0.0333	0.0403	0.0409
	1600	0.0170	0.0205	0.0231	0.0239	0.0245	0.0170	0.0205	0.0233	0.0242	0.0251
	3200	0.0179	0.0177	0.0173	0.0164	0.0170	0.0179	0.0177	0.0173	0.0164	0.0171
	6400	0.0177	0.0166	0.0147	0.0125	0.0130	0.0177	0.0166	0.0147	0.0125	0.0130
1.00	100	0.0259	0.0672	0.0901	0.1155	0.1190	0.0767	0.1846	0.2748	0.3875	0.4502
	200	0.0315	0.0741	0.0849	0.1093	0.1107	0.0528	0.1301	0.1812	0.2352	0.2923
	400	0.0264	0.0494	0.0738	0.0921	0.0962	0.0308	0.0744	0.1035	0.1265	0.1455
	800	0.0194	0.0417	0.0552	0.0622	0.0678	0.0197	0.0454	0.0677	0.0801	0.0821
	1600	0.0168	0.0287	0.0397	0.0447	0.0450	0.0168	0.0288	0.0416	0.0472	0.0491
	3200	0.0182	0.0213	0.0254	0.0294	0.0320	0.0182	0.0213	0.0257	0.0301	0.0327
	6400	0.0178	0.0178	0.0181	0.0196	0.0218	0.0178	0.0178	0.0181	0.0197	0.0223
2.00	100	0.0036	0.0032	0.0024	0.0013	0.0008	0.2859	0.7817	1.0884	1.4798	1.8207
	200	0.0036	0.0032	0.0024	0.0668	0.0008	0.2666	0.7382	1.0145	1.4084	1.8834
	400	0.0121	0.0211	0.0430	0.0920	0.0213	0.2010	0.5514	0.8411	1.1637	1.4428
	800	0.0106	0.0032	0.0024	0.0013	0.0740	0.1610	0.4011	0.7363	1.0270	1.4034
	1600	0.0185	0.0032	0.0024	0.0013	0.0802	0.1077	0.4311	0.5550	0.8538	0.8346
	3200	0.0267	0.0032	0.0024	0.0013	0.0008	0.0550	0.2867	0.4715	0.7489	0.8256
	6400	0.0248	0.0444	0.0024	0.0013	0.0008	0.0298	0.1729	0.3987	0.5710	0.6829

Table E.8: Median Average Square Error for Estimated Marginal Product, with Heteroskedasticity

		SVKZ					HMS				
ρ	n	(1,1)	(1,2)	(2,2)	(2,3)	(3,3)	(1,1)	(1,2)	(2,2)	(2,3)	(3,3)
0.00	100	0.0833	0.0737	0.1278	0.1469	0.1875	0.0836	0.0795	0.1475	0.1673	0.2359
	200	0.0472	0.0406	0.0761	0.0885	0.1434	0.0472	0.0422	0.0768	0.0886	0.1515
	400	0.0277	0.0242	0.0475	0.0527	0.0798	0.0277	0.0242	0.0478	0.0528	0.0814
	800	0.0192	0.0150	0.0350	0.0359	0.0528	0.0192	0.0150	0.0350	0.0359	0.0533
	1600	0.0138	0.0102	0.0189	0.0233	0.0325	0.0138	0.0102	0.0189	0.0233	0.0325
	3200	0.0099	0.0079	0.0115	0.0124	0.0213	0.0099	0.0079	0.0115	0.0124	0.0214
	6400	0.0071	0.0064	0.0083	0.0079	0.0148	0.0071	0.0064	0.0083	0.0079	0.0148
0.25	100	0.0882	0.0748	0.1470	0.1460	0.2188	0.0889	0.0835	0.1703	0.1683	0.2874
	200	0.0542	0.0500	0.0834	0.0997	0.1533	0.0547	0.0505	0.0859	0.1022	0.1633
	400	0.0310	0.0239	0.0586	0.0595	0.0859	0.0310	0.0239	0.0586	0.0605	0.0880
	800	0.0193	0.0156	0.0356	0.0365	0.0552	0.0193	0.0156	0.0356	0.0366	0.0572
	1600	0.0137	0.0103	0.0205	0.0219	0.0401	0.0137	0.0103	0.0205	0.0219	0.0405
	3200	0.0115	0.0086	0.0124	0.0135	0.0232	0.0115	0.0086	0.0124	0.0135	0.0232
	6400	0.0076	0.0070	0.0084	0.0078	0.0162	0.0076	0.0070	0.0084	0.0078	0.0162
0.50	100	0.1069	0.0972	0.2058	0.1847	0.2576	0.1219	0.1319	0.2573	0.2745	0.3936
	200	0.0729	0.0720	0.1199	0.1337	0.1979	0.0733	0.0749	0.1282	0.1546	0.2414
	400	0.0353	0.0359	0.0867	0.0851	0.1273	0.0353	0.0364	0.0880	0.0896	0.1311
	800	0.0206	0.0210	0.0468	0.0481	0.0748	0.0206	0.0210	0.0468	0.0485	0.0757
	1600	0.0153	0.0118	0.0255	0.0306	0.0534	0.0153	0.0118	0.0255	0.0308	0.0542
	3200	0.0132	0.0103	0.0138	0.0185	0.0298	0.0132	0.0103	0.0138	0.0185	0.0298
	6400	0.0079	0.0079	0.0094	0.0102	0.0207	0.0079	0.0079	0.0094	0.0102	0.0209
0.75	100	0.1519	0.1129	0.2258	0.2155	0.3269	0.2234	0.2226	0.4871	0.4396	0.6332
	200	0.1050	0.1365	0.1782	0.1851	0.2508	0.1150	0.1714	0.2775	0.2812	0.4314
	400	0.0516	0.0609	0.1377	0.1296	0.1922	0.0516	0.0679	0.1615	0.1678	0.2543
	800	0.0304	0.0335	0.0799	0.0891	0.1315	0.0306	0.0340	0.0819	0.0963	0.1360
	1600	0.0186	0.0178	0.0433	0.0457	0.0798	0.0186	0.0178	0.0436	0.0467	0.0835
	3200	0.0164	0.0118	0.0208	0.0293	0.0449	0.0164	0.0118	0.0208	0.0294	0.0457
	6400	0.0087	0.0093	0.0119	0.0156	0.0332	0.0087	0.0093	0.0119	0.0156	0.0332
1.00	100	0.1504	0.1649	0.2677	0.2354	0.3579	0.4225	0.3933	0.8651	0.6515	0.8998
	200	0.1708	0.1431	0.2045	0.2275	0.2889	0.2664	0.3430	0.6075	0.5562	0.7256
	400	0.0858	0.0890	0.1941	0.1897	0.2587	0.1026	0.1624	0.3796	0.3205	0.5062
	800	0.0529	0.0726	0.1469	0.1372	0.2200	0.0538	0.0836	0.2025	0.2322	0.2779
	1600	0.0252	0.0314	0.0905	0.0919	0.1386	0.0252	0.0314	0.1001	0.1096	0.1724
	3200	0.0194	0.0156	0.0421	0.0630	0.0840	0.0194	0.0156	0.0425	0.0672	0.0884
	6400	0.0110	0.0118	0.0195	0.0313	0.0646	0.0110	0.0118	0.0195	0.0314	0.0671
2.00	100	0.1810	0.1606	0.4240	0.3151	0.4970	1.5143	0.9216	1.6562	0.8609	0.9649
	200	0.1882	0.1200	0.2492	0.2869	0.3211	1.6679	1.0047	1.2736	0.8969	1.0565
	400	0.1535	0.1006	0.1963	0.1680	0.2475	1.1596	0.9809	1.5572	0.8425	1.1034
	800	0.1119	0.0553	0.1020	0.0986	0.1546	0.9428	0.7703	1.3756	0.8066	1.0385
	1600	0.1050	0.0365	0.0530	0.0516	0.1216	0.5052	0.6653	1.3412	0.7155	1.0097
	3200	0.0937	0.0465	0.0283	0.0217	0.0477	0.1938	0.4422	1.0167	0.8166	1.0893
	6400	0.0566	0.0536	0.0203	0.0104	0.0200	0.0698	0.3163	0.8885	0.6330	0.9004

Table E.9: IMSE for Frontier Estimates with No Heteroskedasticity (Exponential Inefficiency)

		SVKZ					HMS				
ρ	n	(1,1)	(1,2)	(2,2)	(2,3)	(3,3)	(1,1)	(1,2)	(2,2)	(2,3)	(3,3)
0.00	100	0.0569	0.0611	0.0727	0.0750	0.0839	0.0718	0.0925	0.1241	0.1328	0.1473
	200	0.0575	0.0575	0.0623	0.0664	0.0695	0.0645	0.0723	0.0892	0.1043	0.1113
	400	0.0557	0.0517	0.0575	0.0610	0.0663	0.0563	0.0560	0.0693	0.0785	0.0908
	800	0.0600	0.0505	0.0517	0.0585	0.0621	0.0602	0.0519	0.0548	0.0672	0.0727
	1600	0.0640	0.0518	0.0493	0.0523	0.0560	0.0642	0.0519	0.0501	0.0543	0.0588
	3200	0.0652	0.0560	0.0486	0.0498	0.0522	0.0652	0.0560	0.0487	0.0501	0.0527
	6400	0.0671	0.0589	0.0509	0.0482	0.0503	0.0671	0.0589	0.0509	0.0482	0.0504
0.25	100	0.0581	0.0626	0.0748	0.0776	0.0874	0.0736	0.0959	0.1287	0.1388	0.1518
	200	0.0574	0.0587	0.0642	0.0686	0.0719	0.0648	0.0749	0.0930	0.1086	0.1173
	400	0.0556	0.0525	0.0586	0.0615	0.0678	0.0564	0.0574	0.0706	0.0794	0.0946
	800	0.0595	0.0505	0.0523	0.0591	0.0631	0.0597	0.0518	0.0557	0.0684	0.0738
	1600	0.0637	0.0517	0.0493	0.0527	0.0562	0.0639	0.0519	0.0502	0.0546	0.0593
	3200	0.0650	0.0555	0.0482	0.0498	0.0523	0.0650	0.0556	0.0483	0.0501	0.0527
	6400	0.0671	0.0585	0.0503	0.0479	0.0502	0.0671	0.0585	0.0503	0.0480	0.0503
0.50	100	0.0626	0.0692	0.0826	0.0872	0.0981	0.0804	0.1086	0.1466	0.1565	0.1721
	200	0.0588	0.0629	0.0702	0.0749	0.0781	0.0676	0.0820	0.1043	0.1224	0.1336
	400	0.0555	0.0548	0.0617	0.0650	0.0724	0.0566	0.0608	0.0759	0.0862	0.1049
	800	0.0584	0.0506	0.0539	0.0610	0.0655	0.0587	0.0520	0.0583	0.0718	0.0781
	1600	0.0628	0.0508	0.0495	0.0538	0.0573	0.0630	0.0511	0.0506	0.0563	0.0608
	3200	0.0643	0.0541	0.0473	0.0500	0.0526	0.0643	0.0541	0.0474	0.0504	0.0533
	6400	0.0666	0.0572	0.0488	0.0471	0.0498	0.0666	0.0572	0.0489	0.0472	0.0499
0.75	100	0.0730	0.0841	0.0991	0.1042	0.1160	0.0963	0.1392	0.1859	0.1984	0.2202
	200	0.0627	0.0716	0.0802	0.0863	0.0909	0.0741	0.0981	0.1245	0.1509	0.1666
	400	0.0562	0.0591	0.0674	0.0721	0.0802	0.0578	0.0682	0.0860	0.1012	0.1232
	800	0.0572	0.0518	0.0568	0.0652	0.0703	0.0575	0.0539	0.0633	0.0802	0.0883
	1600	0.0613	0.0495	0.0499	0.0556	0.0596	0.0615	0.0500	0.0516	0.0589	0.0649
	3200	0.0631	0.0517	0.0460	0.0501	0.0533	0.0631	0.0517	0.0461	0.0509	0.0546
	6400	0.0658	0.0550	0.0464	0.0459	0.0491	0.0658	0.0550	0.0464	0.0460	0.0492
1.00	100	0.0900	0.1035	0.1216	0.1264	0.1413	0.1343	0.2001	0.2636	0.2753	0.3095
	200	0.0730	0.0852	0.0961	0.1045	0.1090	0.0912	0.1301	0.1686	0.2073	0.2286
	400	0.0590	0.0676	0.0777	0.0847	0.0942	0.0624	0.0838	0.1083	0.1330	0.1632
	800	0.0563	0.0551	0.0621	0.0719	0.0782	0.0568	0.0590	0.0730	0.0954	0.1081
	1600	0.0593	0.0485	0.0517	0.0587	0.0638	0.0595	0.0496	0.0548	0.0646	0.0729
	3200	0.0615	0.0486	0.0446	0.0505	0.0548	0.0615	0.0486	0.0448	0.0522	0.0573
	6400	0.0645	0.0517	0.0433	0.0443	0.0484	0.0645	0.0517	0.0433	0.0449	0.0488
2.00	100	0.2014	0.2256	0.2579	0.2768	0.2989	0.7316	0.9399	1.0986	1.1562	1.2255
	200	0.1632	0.1860	0.2051	0.2239	0.2321	0.4790	0.6829	0.7962	0.8903	0.9529
	400	0.1266	0.1497	0.1696	0.1822	0.1965	0.3126	0.4452	0.5245	0.6322	0.7184
	800	0.0958	0.1206	0.1330	0.1463	0.1569	0.1701	0.2788	0.3438	0.4053	0.4737
	1600	0.0687	0.0887	0.1042	0.1165	0.1254	0.0880	0.1597	0.2115	0.2579	0.3004
	3200	0.0580	0.0615	0.0753	0.0870	0.0950	0.0615	0.0827	0.1169	0.1458	0.1756
	6400	0.0560	0.0426	0.0526	0.0625	0.0715	0.0561	0.0446	0.0630	0.0805	0.0996

Table E.10: IMSE for Frontier Estimates with Heteroskedastic, Exponential Inefficiency and Noise

		SVKZ					HMS				
ρ	n	(1,1)	(1,2)	(2,2)	(2,3)	(3,3)	(1,1)	(1,2)	(2,2)	(2,3)	(3,3)
0.00	100	0.0582	0.0661	0.0687	0.0765	0.0793	0.0776	0.1112	0.1243	0.1392	0.1514
	200	0.0569	0.0623	0.0660	0.0701	0.0731	0.0685	0.0907	0.1020	0.1124	0.1244
	400	0.0578	0.0537	0.0574	0.0633	0.0658	0.0610	0.0639	0.0733	0.0855	0.0920
	800	0.0598	0.0525	0.0541	0.0583	0.0634	0.0610	0.0552	0.0606	0.0670	0.0778
	1600	0.0635	0.0548	0.0500	0.0537	0.0575	0.0636	0.0556	0.0516	0.0561	0.0619
	3200	0.0654	0.0570	0.0502	0.0507	0.0546	0.0655	0.0572	0.0505	0.0513	0.0563
	6400	0.0673	0.0603	0.0517	0.0494	0.0519	0.0673	0.0603	0.0519	0.0495	0.0522
0.25	100	0.0594	0.0680	0.0704	0.0798	0.0840	0.0805	0.1138	0.1279	0.1475	0.1607
	200	0.0572	0.0629	0.0673	0.0715	0.0746	0.0692	0.0915	0.1040	0.1158	0.1298
	400	0.0579	0.0545	0.0589	0.0642	0.0676	0.0612	0.0657	0.0752	0.0882	0.0960
	800	0.0598	0.0527	0.0549	0.0588	0.0642	0.0608	0.0557	0.0620	0.0681	0.0790
	1600	0.0633	0.0544	0.0500	0.0543	0.0578	0.0633	0.0552	0.0517	0.0569	0.0624
	3200	0.0652	0.0565	0.0500	0.0508	0.0549	0.0653	0.0568	0.0503	0.0515	0.0567
	6400	0.0671	0.0599	0.0512	0.0492	0.0518	0.0671	0.0599	0.0514	0.0493	0.0520
0.50	100	0.0640	0.0749	0.0781	0.0907	0.0946	0.0894	0.1253	0.1439	0.1689	0.1837
	200	0.0586	0.0667	0.0725	0.0775	0.0811	0.0715	0.0992	0.1133	0.1283	0.1453
	400	0.0579	0.0568	0.0623	0.0681	0.0719	0.0618	0.0702	0.0817	0.0972	0.1057
	800	0.0593	0.0530	0.0564	0.0608	0.0672	0.0602	0.0567	0.0647	0.0719	0.0852
	1600	0.0625	0.0535	0.0503	0.0555	0.0592	0.0626	0.0545	0.0523	0.0589	0.0645
	3200	0.0644	0.0551	0.0494	0.0509	0.0554	0.0646	0.0554	0.0500	0.0517	0.0576
	6400	0.0663	0.0585	0.0497	0.0484	0.0518	0.0663	0.0585	0.0499	0.0486	0.0521
0.75	100	0.0740	0.0881	0.0943	0.1081	0.1137	0.1074	0.1523	0.1808	0.2176	0.2360
	200	0.0628	0.0750	0.0826	0.0891	0.0932	0.0786	0.1145	0.1350	0.1555	0.1780
	400	0.0584	0.0616	0.0683	0.0753	0.0801	0.0636	0.0786	0.0945	0.1141	0.1258
	800	0.0587	0.0539	0.0595	0.0649	0.0719	0.0597	0.0591	0.0705	0.0805	0.0969
	1600	0.0613	0.0521	0.0509	0.0576	0.0614	0.0614	0.0535	0.0535	0.0626	0.0691
	3200	0.0633	0.0530	0.0484	0.0509	0.0562	0.0634	0.0534	0.0494	0.0522	0.0591
	6400	0.0652	0.0563	0.0472	0.0471	0.0511	0.0652	0.0563	0.0473	0.0473	0.0516
1.00	100	0.0912	0.1087	0.1179	0.1319	0.1404	0.1472	0.2114	0.2549	0.3030	0.3290
	200	0.0723	0.0894	0.0986	0.1074	0.1115	0.0953	0.1474	0.1809	0.2066	0.2445
	400	0.0614	0.0699	0.0791	0.0872	0.0946	0.0693	0.0955	0.1189	0.1456	0.1679
	800	0.0583	0.0567	0.0656	0.0714	0.0797	0.0599	0.0652	0.0823	0.0952	0.1173
	1600	0.0596	0.0514	0.0529	0.0613	0.0656	0.0599	0.0539	0.0573	0.0694	0.0776
	3200	0.0614	0.0502	0.0474	0.0516	0.0576	0.0616	0.0509	0.0492	0.0539	0.0616
	6400	0.0635	0.0531	0.0444	0.0457	0.0507	0.0635	0.0532	0.0446	0.0460	0.0514
2.00	100	0.2025	0.2361	0.2608	0.2781	0.2979	0.7914	0.9428	1.0901	1.1901	1.2334
	200	0.1632	0.1942	0.2133	0.2273	0.2342	0.5299	0.7197	0.8252	0.9139	0.9952
	400	0.1275	0.1599	0.1708	0.1851	0.1977	0.3153	0.4668	0.5763	0.6576	0.7611
	800	0.0961	0.1188	0.1384	0.1478	0.1585	0.1683	0.2940	0.3635	0.4359	0.4883
	1600	0.0718	0.0910	0.1048	0.1205	0.1262	0.1001	0.1712	0.2161	0.2743	0.3094
	3200	0.0582	0.0632	0.0776	0.0871	0.0988	0.0650	0.0915	0.1249	0.1532	0.1800
	6400	0.0542	0.0450	0.0533	0.0633	0.0723	0.0547	0.0510	0.0682	0.0848	0.1005

Table E.11: Median Average Square Error for Derivatives of Expected Efficiency $\mu_\eta(Z)$, with Heteroskedasticity

		SVKZ					HMS				
ρ	n	(1,1)	(1,2)	(2,2)	(2,3)	(3,3)	(1,1)	(1,2)	(2,2)	(2,3)	(3,3)
0.00	100	0.0410	0.0786	0.1088	0.1347	0.1612	0.0411	0.0789	0.1106	0.1393	0.1683
	200	0.0336	0.0702	0.0872	0.1090	0.1251	0.0336	0.0707	0.0879	0.1100	0.1293
	400	0.0299	0.0506	0.0619	0.0768	0.0889	0.0299	0.0509	0.0622	0.0772	0.0891
	800	0.0295	0.0432	0.0499	0.0552	0.0632	0.0295	0.0434	0.0499	0.0553	0.0632
	1600	0.0327	0.0362	0.0399	0.0388	0.0402	0.0327	0.0362	0.0399	0.0388	0.0402
	3200	0.0331	0.0340	0.0304	0.0288	0.0290	0.0331	0.0340	0.0304	0.0288	0.0290
	6400	0.0327	0.0320	0.0288	0.0214	0.0204	0.0327	0.0320	0.0288	0.0214	0.0204
0.25	100	0.0453	0.0812	0.1073	0.1508	0.1722	0.0455	0.0826	0.1110	0.1558	0.1854
	200	0.0347	0.0676	0.0899	0.1126	0.1389	0.0347	0.0688	0.0910	0.1134	0.1414
	400	0.0298	0.0526	0.0632	0.0789	0.0920	0.0298	0.0527	0.0638	0.0790	0.0928
	800	0.0292	0.0435	0.0495	0.0567	0.0664	0.0292	0.0438	0.0495	0.0568	0.0664
	1600	0.0320	0.0376	0.0405	0.0407	0.0418	0.0320	0.0376	0.0405	0.0407	0.0418
	3200	0.0327	0.0341	0.0314	0.0291	0.0289	0.0327	0.0341	0.0314	0.0291	0.0289
	6400	0.0322	0.0323	0.0290	0.0215	0.0206	0.0322	0.0323	0.0290	0.0215	0.0206
0.50	100	0.0485	0.0956	0.1258	0.1870	0.2212	0.0495	0.1006	0.1339	0.2040	0.2460
	200	0.0356	0.0744	0.1011	0.1291	0.1561	0.0357	0.0757	0.1022	0.1315	0.1601
	400	0.0308	0.0552	0.0740	0.0906	0.1046	0.0308	0.0555	0.0743	0.0909	0.1057
	800	0.0283	0.0454	0.0512	0.0613	0.0719	0.0283	0.0455	0.0514	0.0614	0.0719
	1600	0.0323	0.0384	0.0426	0.0431	0.0461	0.0323	0.0384	0.0426	0.0431	0.0462
	3200	0.0329	0.0338	0.0323	0.0310	0.0312	0.0329	0.0338	0.0323	0.0310	0.0313
	6400	0.0323	0.0321	0.0287	0.0227	0.0221	0.0323	0.0321	0.0287	0.0227	0.0221
0.75	100	0.0523	0.1188	0.1633	0.2433	0.2841	0.0628	0.1330	0.1998	0.3044	0.3610
	200	0.0410	0.0990	0.1305	0.1615	0.1970	0.0415	0.1036	0.1329	0.1705	0.2151
	400	0.0324	0.0635	0.0879	0.1068	0.1299	0.0324	0.0637	0.0882	0.1085	0.1315
	800	0.0290	0.0504	0.0597	0.0743	0.0850	0.0290	0.0504	0.0599	0.0744	0.0854
	1600	0.0328	0.0407	0.0454	0.0502	0.0515	0.0328	0.0407	0.0454	0.0502	0.0518
	3200	0.0327	0.0347	0.0348	0.0335	0.0336	0.0327	0.0347	0.0350	0.0335	0.0336
	6400	0.0327	0.0325	0.0294	0.0243	0.0244	0.0327	0.0325	0.0294	0.0243	0.0244
1.00	100	0.0621	0.1488	0.2134	0.2925	0.3332	0.0996	0.2125	0.3257	0.5003	0.6213
	200	0.0528	0.1293	0.1737	0.2251	0.2700	0.0599	0.1454	0.1980	0.2629	0.3237
	400	0.0370	0.0813	0.1207	0.1472	0.1736	0.0382	0.0830	0.1229	0.1539	0.1836
	800	0.0311	0.0596	0.0771	0.0940	0.1086	0.0311	0.0599	0.0777	0.0952	0.1108
	1600	0.0331	0.0443	0.0554	0.0623	0.0658	0.0331	0.0443	0.0555	0.0626	0.0664
	3200	0.0324	0.0366	0.0389	0.0397	0.0405	0.0324	0.0366	0.0390	0.0398	0.0405
	6400	0.0329	0.0338	0.0318	0.0276	0.0284	0.0329	0.0338	0.0318	0.0276	0.0284
2.00	100	0.0306	0.1721	0.2567	0.3573	0.3393	0.4477	1.0962	1.5814	2.4811	2.7370
	200	0.0544	0.2335	0.3046	0.4092	0.5095	0.3087	0.8262	1.3476	1.6593	2.4729
	400	0.0596	0.2030	0.3085	0.4492	0.4146	0.1898	0.5268	0.8158	1.0988	1.5071
	800	0.0649	0.1877	0.2363	0.3406	0.3768	0.1117	0.3497	0.4913	0.6653	0.7444
	1600	0.0534	0.1380	0.1876	0.2703	0.2838	0.0703	0.2070	0.2792	0.3694	0.4194
	3200	0.0420	0.0950	0.1355	0.1739	0.1998	0.0441	0.1121	0.1681	0.2009	0.2323
	6400	0.0372	0.0635	0.0966	0.1135	0.1208	0.0372	0.0649	0.1001	0.1205	0.1331

Table E.12: Median Average Square Error for Estimated Marginal Product, with Heteroskedasticity

		SVKZ					HMS				
ρ	n	(1,1)	(1,2)	(2,2)	(2,3)	(3,3)	(1,1)	(1,2)	(2,2)	(2,3)	(3,3)
0.00	100	0.1611	0.1483	0.2695	0.2616	0.3914	0.1611	0.1510	0.2752	0.2649	0.4131
	200	0.1071	0.1163	0.1858	0.2154	0.2865	0.1071	0.1169	0.1884	0.2163	0.2965
	400	0.0769	0.0853	0.1541	0.1733	0.2503	0.0769	0.0853	0.1547	0.1733	0.2525
	800	0.0576	0.0577	0.1178	0.1199	0.1782	0.0576	0.0579	0.1178	0.1201	0.1782
	1600	0.0418	0.0410	0.0839	0.0781	0.1132	0.0418	0.0410	0.0839	0.0781	0.1132
	3200	0.0421	0.0306	0.0462	0.0577	0.0808	0.0421	0.0306	0.0462	0.0577	0.0808
	6400	0.0336	0.0277	0.0360	0.0362	0.0534	0.0336	0.0277	0.0360	0.0362	0.0534
0.25	100	0.1720	0.1616	0.2657	0.2730	0.3777	0.1720	0.1632	0.2758	0.2793	0.4079
	200	0.1132	0.1196	0.2004	0.2212	0.3178	0.1132	0.1256	0.2022	0.2225	0.3194
	400	0.0745	0.0831	0.1747	0.1707	0.2548	0.0745	0.0833	0.1783	0.1707	0.2659
	800	0.0554	0.0633	0.1148	0.1347	0.1807	0.0554	0.0637	0.1157	0.1369	0.1807
	1600	0.0423	0.0439	0.0821	0.0862	0.1218	0.0423	0.0439	0.0821	0.0862	0.1218
	3200	0.0441	0.0327	0.0477	0.0589	0.0796	0.0441	0.0327	0.0477	0.0589	0.0796
	6400	0.0315	0.0283	0.0369	0.0389	0.0563	0.0315	0.0283	0.0369	0.0389	0.0563
0.50	100	0.2023	0.1821	0.3191	0.3537	0.4435	0.2059	0.1966	0.3532	0.3862	0.5297
	200	0.1247	0.1462	0.2371	0.2225	0.3630	0.1257	0.1482	0.2450	0.2269	0.3914
	400	0.0744	0.0867	0.1882	0.1885	0.2857	0.0744	0.0875	0.1906	0.1916	0.2892
	800	0.0637	0.0711	0.1287	0.1559	0.2191	0.0637	0.0714	0.1288	0.1572	0.2194
	1600	0.0419	0.0461	0.0963	0.0975	0.1298	0.0419	0.0461	0.0963	0.0975	0.1314
	3200	0.0496	0.0358	0.0557	0.0653	0.0869	0.0496	0.0358	0.0557	0.0653	0.0871
	6400	0.0317	0.0292	0.0403	0.0415	0.0591	0.0317	0.0292	0.0403	0.0415	0.0591
0.75	100	0.2555	0.2485	0.4034	0.4266	0.5228	0.2973	0.2952	0.4949	0.5301	0.6250
	200	0.1703	0.1862	0.3299	0.2923	0.4323	0.1732	0.2031	0.3444	0.3131	0.5366
	400	0.0954	0.1027	0.2389	0.2193	0.3383	0.0954	0.1032	0.2400	0.2283	0.3427
	800	0.0736	0.0831	0.1605	0.1794	0.2602	0.0736	0.0831	0.1620	0.1808	0.2639
	1600	0.0497	0.0496	0.1225	0.1151	0.1613	0.0497	0.0496	0.1225	0.1153	0.1635
	3200	0.0527	0.0386	0.0657	0.0760	0.1012	0.0527	0.0386	0.0657	0.0762	0.1012
	6400	0.0330	0.0313	0.0464	0.0472	0.0659	0.0330	0.0313	0.0464	0.0472	0.0659
1.00	100	0.3398	0.3178	0.5147	0.4138	0.5521	0.4786	0.4126	0.8098	0.6485	0.7341
	200	0.2220	0.2615	0.4100	0.3645	0.4973	0.2495	0.3198	0.4983	0.4418	0.6420
	400	0.1319	0.1440	0.3357	0.2917	0.4269	0.1347	0.1463	0.3543	0.3230	0.4521
	800	0.0911	0.1060	0.2223	0.2524	0.3377	0.0911	0.1060	0.2252	0.2641	0.3453
	1600	0.0568	0.0623	0.1486	0.1386	0.1940	0.0568	0.0623	0.1493	0.1400	0.1993
	3200	0.0536	0.0460	0.0840	0.0956	0.1230	0.0536	0.0460	0.0840	0.0956	0.1230
	6400	0.0365	0.0352	0.0572	0.0584	0.0760	0.0365	0.0352	0.0572	0.0584	0.0760
2.00	100	0.3916	0.4025	0.7532	0.4992	0.7055	2.5102	1.2524	1.5276	0.9308	0.9905
	200	0.3673	0.3693	0.5471	0.4675	0.5127	2.1563	1.3826	1.4545	0.8951	0.9614
	400	0.3448	0.2839	0.5841	0.4649	0.6011	1.0959	0.9290	1.3798	0.8451	0.9953
	800	0.3194	0.2910	0.5647	0.4399	0.5710	0.6137	0.8006	1.2553	0.9122	0.9232
	1600	0.2308	0.2521	0.4977	0.4450	0.5550	0.3051	0.4315	0.9295	0.6785	0.8654
	3200	0.1506	0.2004	0.3797	0.3834	0.5072	0.1600	0.2524	0.5303	0.5031	0.6552
	6400	0.0892	0.0943	0.2753	0.2419	0.3257	0.0892	0.0992	0.3090	0.2808	0.4229

Table E.13: Differences in IMSE for Frontier Estimates, Exponential versus Half-Normal Inefficiency, No Heteroskedasticity

		SVKZ					HMS				
ρ	n	(1,1)	(1,2)	(2,2)	(2,3)	(3,3)	(1,1)	(1,2)	(2,2)	(2,3)	(3,3)
0.00	100	0.0376	0.0241	0.0266	0.0202	0.0208	0.0558	0.0652	0.0898	0.0940	0.1035
	200	0.0481	0.0347	0.0335	0.0327	0.0294	0.0563	0.0537	0.0664	0.0780	0.0821
	400	0.0513	0.0392	0.0399	0.0405	0.0424	0.0521	0.0447	0.0540	0.0615	0.0716
	800	0.0579	0.0438	0.0402	0.0454	0.0481	0.0581	0.0453	0.0438	0.0551	0.0603
	1600	0.0630	0.0481	0.0416	0.0429	0.0462	0.0631	0.0483	0.0425	0.0451	0.0494
	3200	0.0646	0.0539	0.0434	0.0424	0.0446	0.0646	0.0540	0.0435	0.0428	0.0452
	6400	0.0668	0.0578	0.0475	0.0422	0.0438	0.0668	0.0578	0.0475	0.0423	0.0439
0.25	100	0.0360	0.0219	0.0250	0.0197	0.0205	0.0562	0.0668	0.0920	0.0979	0.1056
	200	0.0471	0.0336	0.0320	0.0320	0.0293	0.0558	0.0550	0.0686	0.0807	0.0867
	400	0.0506	0.0383	0.0393	0.0391	0.0415	0.0517	0.0449	0.0542	0.0614	0.0741
	800	0.0571	0.0432	0.0396	0.0448	0.0482	0.0573	0.0446	0.0438	0.0555	0.0607
	1600	0.0626	0.0477	0.0409	0.0426	0.0458	0.0627	0.0479	0.0418	0.0448	0.0493
	3200	0.0644	0.0533	0.0426	0.0420	0.0442	0.0644	0.0533	0.0427	0.0424	0.0447
	6400	0.0668	0.0573	0.0467	0.0416	0.0432	0.0668	0.0573	0.0467	0.0416	0.0433
0.50	100	0.0303	0.0177	0.0225	0.0182	0.0205	0.0588	0.0740	0.1028	0.1057	0.1163
	200	0.0438	0.0289	0.0282	0.0289	0.0266	0.0558	0.0579	0.0750	0.0896	0.0971
	400	0.0485	0.0344	0.0356	0.0357	0.0387	0.0503	0.0450	0.0561	0.0648	0.0799
	800	0.0553	0.0406	0.0368	0.0425	0.0460	0.0556	0.0426	0.0436	0.0563	0.0623
	1600	0.0612	0.0455	0.0385	0.0409	0.0440	0.0614	0.0459	0.0401	0.0441	0.0489
	3200	0.0635	0.0512	0.0401	0.0403	0.0427	0.0635	0.0512	0.0402	0.0409	0.0436
	6400	0.0662	0.0556	0.0441	0.0393	0.0415	0.0662	0.0556	0.0442	0.0394	0.0416
0.75	100	0.0239	0.0169	0.0236	0.0204	0.0258	0.0648	0.0890	0.1223	0.1238	0.1394
	200	0.0370	0.0209	0.0227	0.0256	0.0246	0.0563	0.0653	0.0837	0.1037	0.1129
	400	0.0434	0.0255	0.0278	0.0290	0.0336	0.0481	0.0467	0.0594	0.0716	0.0893
	800	0.0519	0.0338	0.0287	0.0359	0.0403	0.0526	0.0402	0.0435	0.0592	0.0666
	1600	0.0589	0.0407	0.0319	0.0351	0.0388	0.0592	0.0419	0.0364	0.0426	0.0488
	3200	0.0619	0.0472	0.0348	0.0356	0.0384	0.0619	0.0472	0.0353	0.0375	0.0413
	6400	0.0652	0.0525	0.0393	0.0346	0.0374	0.0652	0.0525	0.0394	0.0349	0.0380
1.00	100	0.0225	0.0207	0.0320	0.0278	0.0359	0.0772	0.1145	0.1579	0.1541	0.1776
	200	0.0274	0.0161	0.0220	0.0275	0.0271	0.0609	0.0753	0.1000	0.1272	0.1403
	400	0.0321	0.0147	0.0210	0.0245	0.0307	0.0458	0.0501	0.0675	0.0842	0.1068
	800	0.0446	0.0194	0.0154	0.0247	0.0319	0.0480	0.0376	0.0444	0.0634	0.0739
	1600	0.0546	0.0293	0.0166	0.0229	0.0286	0.0553	0.0355	0.0327	0.0414	0.0491
	3200	0.0594	0.0396	0.0219	0.0236	0.0275	0.0594	0.0405	0.0270	0.0326	0.0381
	6400	0.0635	0.0474	0.0304	0.0243	0.0282	0.0635	0.0474	0.0310	0.0274	0.0320
2.00	100	0.0752	0.0840	0.1037	0.1077	0.1206	0.3209	0.4514	0.5414	0.5676	0.5971
	200	0.0554	0.0643	0.0734	0.0871	0.0886	0.1888	0.2968	0.3696	0.4172	0.4606
	400	0.0325	0.0411	0.0589	0.0658	0.0757	0.1146	0.1801	0.2339	0.2840	0.3404
	800	0.0147	0.0174	0.0285	0.0402	0.0517	0.0506	0.0773	0.1144	0.1511	0.1992
	1600	0.0017	−0.0143	0.0001	0.0188	0.0313	0.0219	0.0124	0.0393	0.0851	0.1106
	3200	0.0149	−0.0363	−0.0332	−0.0120	0.0006	0.0315	−0.0132	−0.0133	0.0159	0.0369
	6400	0.0333	−0.0394	−0.0642	−0.0468	−0.0252	0.0427	−0.0093	−0.0388	−0.0246	0.0011

Table E.14: Differences in IMSE for Frontier Estimates, Exponential versus Half-Normal Inefficiency, Heteroskedastic Inefficiency and Noise

		SVKZ					HMS				
ρ	n	(1,1)	(1,2)	(2,2)	(2,3)	(3,3)	(1,1)	(1,2)	(2,2)	(2,3)	(3,3)
0.00	100	0.0376	0.0305	0.0226	0.0219	0.0138	0.0598	0.0814	0.0893	0.0979	0.1070
	200	0.0470	0.0394	0.0368	0.0368	0.0331	0.0596	0.0711	0.0775	0.0861	0.0942
	400	0.0528	0.0402	0.0390	0.0426	0.0422	0.0564	0.0521	0.0571	0.0679	0.0728
	800	0.0576	0.0454	0.0421	0.0449	0.0490	0.0587	0.0485	0.0494	0.0547	0.0650
	1600	0.0623	0.0510	0.0421	0.0443	0.0478	0.0624	0.0519	0.0440	0.0471	0.0526
	3200	0.0648	0.0549	0.0451	0.0436	0.0472	0.0650	0.0551	0.0454	0.0442	0.0490
	6400	0.0671	0.0592	0.0485	0.0436	0.0456	0.0671	0.0592	0.0486	0.0438	0.0459
0.25	100	0.0364	0.0288	0.0216	0.0225	0.0147	0.0612	0.0826	0.0913	0.1045	0.1137
	200	0.0461	0.0374	0.0358	0.0351	0.0312	0.0592	0.0704	0.0786	0.0879	0.0975
	400	0.0523	0.0397	0.0390	0.0413	0.0420	0.0562	0.0528	0.0581	0.0693	0.0754
	800	0.0572	0.0447	0.0418	0.0441	0.0488	0.0583	0.0483	0.0500	0.0550	0.0655
	1600	0.0620	0.0502	0.0414	0.0442	0.0475	0.0620	0.0512	0.0434	0.0473	0.0527
	3200	0.0645	0.0542	0.0445	0.0432	0.0470	0.0646	0.0545	0.0448	0.0439	0.0489
	6400	0.0667	0.0586	0.0477	0.0431	0.0451	0.0667	0.0586	0.0478	0.0432	0.0453
0.50	100	0.0318	0.0246	0.0190	0.0238	0.0157	0.0656	0.0883	0.0995	0.1166	0.1255
	200	0.0428	0.0334	0.0319	0.0317	0.0274	0.0586	0.0734	0.0832	0.0946	0.1055
	400	0.0501	0.0365	0.0358	0.0379	0.0387	0.0551	0.0537	0.0608	0.0738	0.0801
	800	0.0558	0.0422	0.0393	0.0413	0.0469	0.0569	0.0471	0.0500	0.0558	0.0684
	1600	0.0608	0.0478	0.0391	0.0427	0.0460	0.0609	0.0491	0.0418	0.0472	0.0526
	3200	0.0636	0.0521	0.0423	0.0414	0.0456	0.0637	0.0524	0.0430	0.0424	0.0481
	6400	0.0659	0.0569	0.0451	0.0409	0.0437	0.0659	0.0569	0.0453	0.0411	0.0441
0.75	100	0.0254	0.0224	0.0194	0.0271	0.0208	0.0739	0.0993	0.1162	0.1404	0.1502
	200	0.0360	0.0265	0.0269	0.0282	0.0250	0.0592	0.0790	0.0929	0.1074	0.1215
	400	0.0449	0.0285	0.0292	0.0314	0.0326	0.0529	0.0553	0.0660	0.0806	0.0876
	800	0.0527	0.0355	0.0326	0.0344	0.0415	0.0545	0.0446	0.0503	0.0587	0.0739
	1600	0.0584	0.0426	0.0329	0.0376	0.0408	0.0587	0.0451	0.0384	0.0466	0.0529
	3200	0.0620	0.0481	0.0371	0.0368	0.0416	0.0621	0.0487	0.0388	0.0393	0.0460
	6400	0.0645	0.0537	0.0401	0.0362	0.0396	0.0645	0.0538	0.0403	0.0367	0.0406
1.00	100	0.0244	0.0269	0.0278	0.0352	0.0332	0.0893	0.1236	0.1506	0.1795	0.1936
	200	0.0273	0.0223	0.0254	0.0302	0.0275	0.0612	0.0888	0.1102	0.1267	0.1512
	400	0.0343	0.0181	0.0225	0.0262	0.0309	0.0511	0.0596	0.0745	0.0910	0.1066
	800	0.0457	0.0220	0.0203	0.0240	0.0332	0.0502	0.0422	0.0525	0.0620	0.0812
	1600	0.0539	0.0314	0.0186	0.0260	0.0310	0.0550	0.0392	0.0347	0.0464	0.0542
	3200	0.0591	0.0408	0.0252	0.0254	0.0316	0.0593	0.0426	0.0321	0.0348	0.0428
	6400	0.0624	0.0485	0.0312	0.0261	0.0311	0.0624	0.0488	0.0326	0.0293	0.0354
2.00	100	0.0756	0.0937	0.1046	0.1112	0.1157	0.3810	0.4464	0.5450	0.5904	0.6087
	200	0.0566	0.0690	0.0832	0.0916	0.0875	0.2324	0.3264	0.3800	0.4452	0.4710
	400	0.0345	0.0505	0.0589	0.0672	0.0754	0.1195	0.1906	0.2578	0.3067	0.3658
	800	0.0137	0.0155	0.0327	0.0424	0.0530	0.0406	0.0914	0.1305	0.1739	0.2028
	1600	0.0041	−0.0109	0.0005	0.0237	0.0314	0.0277	0.0219	0.0448	0.0926	0.1199
	3200	0.0125	−0.0322	−0.0296	−0.0120	0.0055	0.0304	−0.0036	−0.0075	0.0211	0.0408
	6400	0.0285	−0.0341	−0.0593	−0.0440	−0.0211	0.0386	−0.0035	−0.0331	−0.0200	0.0028

Table E.15: Differences in Median Average Square Error for Derivatives of Expected Efficiency $\mu_\eta(Z)$, Exponential versus Half-Normal Inefficiency, with Heteroskedasticity

		SVKZ					HMS				
ρ	n	(1,1)	(1,2)	(2,2)	(2,3)	(3,3)	(1,1)	(1,2)	(2,2)	(2,3)	(3,3)
0.00	100	0.0228	0.0455	0.0629	0.0776	0.0950	0.0228	0.0439	0.0598	0.0750	0.0899
	200	0.0166	0.0418	0.0550	0.0725	0.0848	0.0166	0.0419	0.0554	0.0735	0.0875
	400	0.0136	0.0287	0.0396	0.0538	0.0633	0.0136	0.0290	0.0398	0.0543	0.0632
	800	0.0119	0.0247	0.0315	0.0384	0.0465	0.0119	0.0249	0.0315	0.0385	0.0464
	1600	0.0150	0.0186	0.0243	0.0264	0.0281	0.0150	0.0186	0.0243	0.0264	0.0281
	3200	0.0152	0.0176	0.0170	0.0188	0.0202	0.0152	0.0176	0.0170	0.0188	0.0202
	6400	0.0152	0.0156	0.0159	0.0127	0.0127	0.0152	0.0156	0.0159	0.0127	0.0127
0.25	100	0.0265	0.0418	0.0571	0.0884	0.1011	0.0258	0.0413	0.0545	0.0831	0.0900
	200	0.0168	0.0372	0.0545	0.0732	0.0922	0.0168	0.0381	0.0547	0.0729	0.0922
	400	0.0130	0.0297	0.0386	0.0533	0.0628	0.0130	0.0298	0.0393	0.0531	0.0633
	800	0.0116	0.0237	0.0304	0.0385	0.0483	0.0116	0.0240	0.0304	0.0385	0.0483
	1600	0.0148	0.0202	0.0246	0.0275	0.0287	0.0148	0.0202	0.0246	0.0275	0.0287
	3200	0.0149	0.0176	0.0177	0.0186	0.0197	0.0149	0.0176	0.0177	0.0186	0.0197
	6400	0.0145	0.0159	0.0160	0.0126	0.0125	0.0145	0.0159	0.0160	0.0126	0.0125
0.50	100	0.0266	0.0444	0.0600	0.1020	0.1369	0.0231	0.0371	0.0488	0.0911	0.0986
	200	0.0160	0.0344	0.0548	0.0746	0.0899	0.0157	0.0343	0.0528	0.0716	0.0837
	400	0.0130	0.0282	0.0408	0.0558	0.0647	0.0130	0.0281	0.0405	0.0549	0.0654
	800	0.0106	0.0236	0.0285	0.0371	0.0478	0.0106	0.0237	0.0286	0.0372	0.0476
	1600	0.0150	0.0203	0.0249	0.0272	0.0301	0.0150	0.0203	0.0249	0.0272	0.0302
	3200	0.0150	0.0168	0.0178	0.0186	0.0198	0.0150	0.0168	0.0179	0.0186	0.0199
	6400	0.0147	0.0158	0.0152	0.0129	0.0126	0.0147	0.0158	0.0152	0.0129	0.0126
0.75	100	0.0255	0.0615	0.0805	0.1398	0.1752	0.0165	0.0284	0.0515	0.1025	0.0781
	200	0.0146	0.0401	0.0661	0.0789	0.1042	0.0131	0.0305	0.0428	0.0650	0.0801
	400	0.0134	0.0272	0.0369	0.0510	0.0668	0.0132	0.0246	0.0324	0.0452	0.0589
	800	0.0110	0.0232	0.0273	0.0352	0.0447	0.0110	0.0231	0.0265	0.0341	0.0446
	1600	0.0158	0.0201	0.0223	0.0262	0.0270	0.0158	0.0201	0.0222	0.0260	0.0266
	3200	0.0148	0.0169	0.0175	0.0171	0.0166	0.0148	0.0169	0.0177	0.0170	0.0165
	6400	0.0150	0.0159	0.0148	0.0118	0.0114	0.0150	0.0159	0.0148	0.0118	0.0114
1.00	100	0.0362	0.0816	0.1233	0.1770	0.2143	0.0229	0.0278	0.0509	0.1128	0.1711
	200	0.0213	0.0551	0.0888	0.1158	0.1593	0.0071	0.0152	0.0168	0.0277	0.0315
	400	0.0106	0.0319	0.0469	0.0551	0.0774	0.0074	0.0086	0.0194	0.0274	0.0381
	800	0.0116	0.0179	0.0219	0.0318	0.0408	0.0113	0.0146	0.0099	0.0151	0.0288
	1600	0.0163	0.0156	0.0157	0.0176	0.0209	0.0163	0.0155	0.0140	0.0154	0.0173
	3200	0.0142	0.0154	0.0134	0.0103	0.0086	0.0142	0.0154	0.0134	0.0097	0.0079
	6400	0.0151	0.0160	0.0137	0.0080	0.0066	0.0151	0.0160	0.0137	0.0079	0.0061
2.00	100	0.0270	0.1689	0.2543	0.3560	0.3384	0.1617	0.3145	0.4930	1.0013	0.9163
	200	0.0508	0.2303	0.3022	0.3424	0.5087	0.0421	0.0880	0.3331	0.2509	0.5895
	400	0.0475	0.1819	0.2655	0.3572	0.3933	−0.0112	−0.0246	−0.0253	−0.0649	0.0642
	800	0.0543	0.1845	0.2339	0.3393	0.3028	−0.0492	−0.0514	−0.2450	−0.3617	−0.6590
	1600	0.0349	0.1348	0.1852	0.2690	0.2036	−0.0375	−0.2241	−0.2758	−0.4844	−0.4152
	3200	0.0154	0.0917	0.1332	0.1726	0.1990	−0.0108	−0.1745	−0.3034	−0.5480	−0.5934
	6400	0.0124	0.0191	0.0943	0.1122	0.1200	0.0074	−0.1080	−0.2986	−0.4505	−0.5498

Table E.16: Differences in Median Average Square Error for Estimated Marginal Product, Exponential versus Half-Normal Inefficiency, with Heteroskedasticity

		SVKZ					HMS				
ρ	n	(1,1)	(1,2)	(2,2)	(2,3)	(3,3)	(1,1)	(1,2)	(2,2)	(2,3)	(3,3)
0.00	100	−0.0423	0.0049	−0.0190	−0.0121	−0.0263	−0.0425	−0.0007	−0.0369	−0.0281	−0.0676
	200	−0.0136	0.0297	0.0111	0.0205	−0.0183	−0.0136	0.0285	0.0111	0.0214	−0.0222
	400	0.0022	0.0264	0.0144	0.0240	0.0091	0.0022	0.0267	0.0144	0.0244	0.0077
	800	0.0103	0.0282	0.0148	0.0193	0.0103	0.0103	0.0284	0.0148	0.0194	0.0099
	1600	0.0189	0.0260	0.0210	0.0155	0.0077	0.0189	0.0260	0.0210	0.0155	0.0077
	3200	0.0233	0.0261	0.0189	0.0164	0.0077	0.0233	0.0261	0.0189	0.0164	0.0076
	6400	0.0256	0.0256	0.0205	0.0134	0.0056	0.0256	0.0256	0.0205	0.0134	0.0056
0.25	100	−0.0430	0.0064	−0.0397	0.0048	−0.0467	−0.0434	−0.0009	−0.0593	−0.0125	−0.1020
	200	−0.0195	0.0176	0.0065	0.0129	−0.0144	−0.0199	0.0183	0.0051	0.0112	−0.0219
	400	−0.0012	0.0286	0.0046	0.0195	0.0060	−0.0012	0.0287	0.0053	0.0185	0.0048
	800	0.0099	0.0278	0.0138	0.0202	0.0112	0.0099	0.0281	0.0139	0.0202	0.0092
	1600	0.0183	0.0273	0.0200	0.0188	0.0016	0.0183	0.0273	0.0200	0.0188	0.0012
	3200	0.0212	0.0254	0.0190	0.0156	0.0057	0.0212	0.0254	0.0190	0.0156	0.0057
	6400	0.0246	0.0252	0.0207	0.0137	0.0043	0.0246	0.0252	0.0207	0.0137	0.0043
0.50	100	−0.0585	−0.0016	−0.0800	0.0023	−0.0364	−0.0724	−0.0313	−0.1234	−0.0705	−0.1476
	200	−0.0373	0.0024	−0.0188	−0.0046	−0.0418	−0.0376	0.0008	−0.0260	−0.0230	−0.0813
	400	−0.0045	0.0193	−0.0128	0.0055	−0.0227	−0.0045	0.0191	−0.0137	0.0013	−0.0254
	800	0.0077	0.0244	0.0044	0.0132	−0.0029	0.0077	0.0245	0.0045	0.0129	−0.0038
	1600	0.0170	0.0266	0.0171	0.0125	−0.0073	0.0170	0.0266	0.0171	0.0123	−0.0080
	3200	0.0197	0.0235	0.0185	0.0125	0.0014	0.0197	0.0235	0.0185	0.0125	0.0016
	6400	0.0245	0.0243	0.0193	0.0125	0.0014	0.0245	0.0243	0.0193	0.0125	0.0012
0.75	100	−0.0996	0.0059	−0.0625	0.0278	−0.0429	−0.1606	−0.0896	−0.2872	−0.1351	−0.2722
	200	−0.0640	−0.0375	−0.0477	−0.0237	−0.0537	−0.0735	−0.0679	−0.1446	−0.1107	−0.2162
	400	−0.0192	0.0026	−0.0498	−0.0227	−0.0622	−0.0192	−0.0042	−0.0733	−0.0593	−0.1229
	800	−0.0014	0.0169	−0.0202	−0.0148	−0.0465	−0.0016	0.0164	−0.0221	−0.0219	−0.0506
	1600	0.0142	0.0228	0.0021	0.0044	−0.0282	0.0142	0.0228	0.0018	0.0036	−0.0317
	3200	0.0163	0.0229	0.0141	0.0042	−0.0113	0.0163	0.0229	0.0143	0.0041	−0.0121
	6400	0.0240	0.0232	0.0176	0.0087	−0.0088	0.0240	0.0232	0.0176	0.0087	−0.0088
1.00	100	−0.0883	−0.0161	−0.0543	0.0571	−0.0246	−0.3229	−0.1808	−0.5394	−0.1512	−0.2786
	200	−0.1180	−0.0139	−0.0308	−0.0024	−0.0189	−0.2065	−0.1977	−0.4095	−0.2933	−0.4018
	400	−0.0488	−0.0077	−0.0734	−0.0425	−0.0851	−0.0644	−0.0795	−0.2567	−0.1666	−0.3226
	800	−0.0218	−0.0130	−0.0698	−0.0432	−0.1114	−0.0228	−0.0236	−0.1249	−0.1370	−0.1671
	1600	0.0079	0.0129	−0.0350	−0.0295	−0.0728	0.0079	0.0129	−0.0445	−0.0470	−0.1060
	3200	0.0130	0.0210	−0.0033	−0.0233	−0.0435	0.0130	0.0210	−0.0035	−0.0273	−0.0479
	6400	0.0218	0.0220	0.0123	−0.0037	−0.0362	0.0218	0.0220	0.0123	−0.0038	−0.0387
2.00	100	−0.1503	0.0115	−0.1673	0.0422	−0.1577	−1.0666	0.1746	−0.0748	1.6202	1.7721
	200	−0.1337	0.1135	0.0553	0.1223	0.1884	−1.3591	−0.1784	0.0740	0.7624	1.4163
	400	−0.0939	0.1024	0.1122	0.2812	0.1672	−0.9697	−0.4541	−0.7414	0.2563	0.4037
	800	−0.0471	0.1324	0.1343	0.2420	0.2221	−0.8310	−0.4206	−0.8843	−0.1412	−0.2941
	1600	−0.0515	0.1015	0.1346	0.2187	0.1622	−0.4349	−0.4583	−1.0620	−0.3462	−0.5904
	3200	−0.0516	0.0484	0.1072	0.1522	0.1521	−0.1497	−0.3301	−0.8486	−0.6157	−0.8571
	6400	−0.0195	0.0099	0.0764	0.1031	0.1008	−0.0326	−0.2514	−0.7884	−0.5125	−0.7674

F Additional Details on Empirical Application

As noted in Section 4, summary statistics for the $p = 5$ input variables X_j, $j \in \{1, 2, 3, 4, 5\}$ and $q = 5$ output variables Y_j, $j \in \{1, 2, 3\}$ used in our empirical illustration appear below in Table F.17. As expected with banking data, the marginal distributions in the original (x, y)-space show evidence of substantial skewness to the right. Also shown in Table F.17 are summary statistics for the reduced-dimensional inputs X_1^*, X_2^* as well as the reduced output variable Y^*. The last three rows of Table F.17 show summary statistics for Z and U after applying the rotation matrix R_d to the reduced dimensional data (X_1^*, X_2^*, Y^*) as described in Section 4.

Table F.18 gives bandwidths optimized via leave-one-out cross validation for the regression of U on Z as well as the regression of squared and cubed residuals from this regression on Z. Table F.18 also shows bounds placed on the bandwidths as discussed in Section 4.

Table F.17: Summary Statistics for Data Used in Empirical Example

	Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum
X_1	5.5235×10^2	6.2096×10^4	1.2591×10^5	1.2217×10^6	2.8082×10^5	1.2806×10^9
X_2	2.0000	1.9000×10^1	3.7000×10^1	2.8455×10^2	8.0500×10^1	2.3337×10^5
X_3	1.0214	8.1779×10^2	2.4850×10^3	1.9554×10^4	6.3354×10^3	2.0447×10^7
X_4	1.3320×10^2	8.0627×10^3	1.5992×10^4	1.9636×10^5	3.4582×10^4	2.2770×10^8
X_5	0.0000	8.5322×10^2	2.4735×10^3	3.9678×10^4	6.8738×10^3	6.1932×10^7
Y_1	4.0940	4.3084×10^4	9.5794×10^4	1.0552×10^6	2.2520×10^5	9.0264×10^8
Y_2	5.6348×10^2	2.3868×10^4	4.8215×10^4	8.6145×10^5	1.0613×10^5	1.3710×10^9
Y_3	1.9573	3.2270×10^2	8.6126×10^2	2.9841×10^4	2.5207×10^3	4.2647×10^7
X_1^*	5.7246×10^{-5}	1.9141×10^{-3}	4.0090×10^{-3}	5.1073×10^{-2}	9.1977×10^{-3}	6.0292×10^1
X_2^*	2.3427×10^{-4}	3.8198×10^{-3}	8.2111×10^{-3}	6.5013×10^{-2}	1.8547×10^{-2}	6.0027×10^1
Y^*	9.8704×10^{-5}	1.5987×10^{-3}	3.3061×10^{-3}	5.1311×10^{-2}	7.7043×10^{-3}	6.2297×10^1
Z_1	5.7246×10^{-5}	1.9141×10^{-3}	4.0090×10^{-3}	5.1073×10^{-2}	9.1977×10^{-3}	6.0292×10^1
Z_2	-3.6212×10^1	-8.0919×10^{-4}	-1.2085×10^{-4}	-2.3315×10^{-2}	4.6720×10^{-4}	1.1536×10^{-1}
U	2.9776×10^{-2}	4.7271×10^{-1}	1.0060	8.9783	2.2713	8.9193×10^3

Table F.18: Bandwidths used for Local-Linear Regressions in Banking Application

	bandwidth	Lower Bound	Upper Bound
Regression of U on Z:			
h_1	3.3305×10^{-1}	1.8638×10^{-2}	3.0583×10^{1}
h_2	2.1973×10^{1}	1.8638×10^{-2}	3.8134×10^{2}
h_3	6.8262×10^{-3}	0.0000	1.0000
h_4	5.9484×10^{-4}	0.0000	1.0000
Regression of Squared Residuals on Z:			
h_1	7.8717×10^{-1}	1.8638×10^{-2}	3.0583×10^{1}
h_2	3.8134×10^{2}	1.8638×10^{-2}	3.8134×10^{2}
h_3	4.4494×10^{-1}	0.0000	1.0000
h_4	1.0000	0.0000	1.0000
Regression of Cubed Residuals on Z:			
h_1	8.9575×10^{-1}	1.8638×10^{-2}	3.0583×10^{1}
h_2	2.0897×10^{2}	1.8638×10^{-2}	3.8134×10^{2}
h_3	1.0552×10^{-1}	0.0000	1.0000
h_4	1.0000	0.0000	1.0000

References

Daouia, A., J. P. Florens, and L. Simar (2020), Robust frontier estimation from noisy data: A tikhonov regularization approach, *Econometrics and Statistics* 14, 1–23.

Fan, J. and I. Gijbels (1996), *Local Polynomial Modelling and Its Applications*, London: Chapman and Hall.

Färe, R., S. Grosskopf, and C. A. K. Lovell (1985), *The Measurement of Efficiency of Production*, Boston: Kluwer-Nijhoff Publishing.

Farrell, M. J. (1957), The measurement of productive efficiency, *Journal of the Royal Statistical Society A* 120, 253–281.

Florens, J. P., L. Simar, and I. Van Keilegom (2020), Estimation of the boundary of a variable observed with symmetric error, *Journal of the American Statistical Association* 115, 425–441.

Hall, P. and L. Simar (2002), Estimating a changepoint, boundary or frontier in the presence of observation error, *Journal of the American Statistical Association* 97, 523–534.

Kneip, A., L. Simar, and I. Van Keilegom (2015), Frontier estimation in the presence of measurement error with unknown variance, *Journal of Econometrics* 184, 379–393.

Kumbhakar, S. C., B. U. Park, L. Simar, and E. G. Tsionas (2007), Nonparametric stochastic frontiers: A local likelihood approach, *Journal of Econometrics* 137, 1–27.

Lang, S. (1964), *A Second Course in Calculus*, Reading, MA: Addison-Wesley Publishing Company, Inc.

Liew, C. K. (1976), Inequality constrained least-squares estimation, *Journal of the American Statistical Association* 71, 746–751.

Marsaglia, G. (1972), Choosing a point from the surface of a sphere, *Annals of Mathematical Statistics* 43, 645–646.

Muller, M. E. (1959), A note on a method for generating points uniformly on n-dimensional spheres, *Communications of the Association for Computing Machinery* 2, 19–20.

Nelder, J. A. and R. Mead (1965), A simplex method for function minimization, *The Computer Journal* 7, 308–313.

Robinson, P. M. (1988), Root-n-consistent semiparametric regression, *Econometrica* 56, 931–954.

Simar, L. and P. W. Wilson (2007), Estimation and inference in two-stage, semi-parametric models of productive efficiency, *Journal of Econometrics* 136, 31–64.

Simar, L. and V. Zelenyuk (2011), Stochastic FDH/DEA estimators for frontier analysis, *Journal of Productivity Analysis* 36, 1–20.

Ziegelmann, F. (2002), Nonparametric estimation of volatility functions: The local exponential estimator, *Econometric Theory* 18, 985–991.