

IZA DP No. 10076

**The Three I's of Public Schools:
Irrelevant Inputs, Insufficient Resources
and Inefficiency**

Daniel J. Henderson
Léopold Simar
Le Wang

July 2016

The Three I's of Public Schools: Irrelevant Inputs, Insufficient Resources and Inefficiency

Daniel J. Henderson
University of Alabama and IZA

Léopold Simar
Université catholique de Louvain

Le Wang
University of Alabama and IZA

Discussion Paper No. 10076
July 2016

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

The Three I's of Public Schools: Irrelevant Inputs, Insufficient Resources and Inefficiency*

We examine the educational production function and efficiency of public school districts in Illinois. Using nonparametric kernel methods, we find that most traditional schooling inputs are irrelevant in determining test scores (even in a very general setting). Property tax caps are the only relevant factor that is related to districts' financial constraints and have predominantly negative associations with test scores. Therefore, insufficient resources may be partially responsible for the lack of growth in test scores. For most other *relevant* inputs, we find substantial heterogeneity in the returns, which helps reconcile some of the puzzling results in the literature. We further find that there exist inefficiencies in school districts. Moreover, the level of test scores, commonly used as a measure of school effectiveness, (while related) differs substantially from our efficiency scores, and standard parametric approaches drastically underestimate school efficiency. We discuss the policy implications of our results.

JEL Classification: C14, C44, I21

Keywords: educational production function, irrelevant inputs, nonparametric kernel, panel data, property tax caps, school inputs, stochastic frontier analysis, student achievement, technical efficiency

Corresponding author:

Le Wang
Department of Economics, Finance and Legal Studies
University of Alabama
Tuscaloosa, AL 35487-0224
USA
E-mail: le.wang.econ@gmail.edu

* The authors would like to thank two anonymous referees for their helpful comments, and Daniel Millimet for sharing the data used in the paper.

1 Introduction

The lack of substantial gains in test scores in the U.S., which have been accompanied by huge increases in expenditures over the last four decades, have caused many to wonder how well schools are using these resources. Such concerns have led to educational policy at both the federal and state levels that emphasize accountability of schools based on measured student (test) performance (Conroy and Arguea, 2008; Hanushek and Raymond, 2005). For example, in the U.S., school districts release information on student learning outcomes to parents and taxpayers as a means to generate pressure for school improvement. Some states and districts even apply sanctions to schools that perform poorly in terms of these measures (Ladd and Walsh, 2002). Twenty-four states have enacted policies allowing for state takeover of a school system. For example, the New Jersey Department of Education has put four school districts deemed inefficient under long-term state control, with the most recent case being the Camden City School District in March 2013.¹

Despite the urgent need for such action, measuring school effectiveness is not a trivial task. The majority of these measures of effectiveness tend to emphasize educational outcomes (e.g., test scores). However, such measures do not necessarily provide a proper evaluation of school efficiency due to their failure to take into account differences in resources available and environmental factors where these schools operate (Ladd and Walsh, 2002). Without taking into account such differences, we may mistakenly attribute the lack of growth in the observed educational outcomes to ineffectiveness, while the real cause may be insufficient resources. Similarly, we may unnecessarily allocate more resources, adding more to the already large size of public budgets, to school districts who are not efficiently utilizing their resources. Distinguishing between insufficient resources and inefficiency in the educational process is crucial for policy debate. For example, as noted in McCarty and Yaisawarng (1993), the State of New Jersey's response to the challenge to its public education financing system (*Abbott v. Burke*) was indeed to argue that "deficiencies in education are primarily related to mismanagement in certain districts rather than to expenditures per pupil".

Thus it is of paramount importance to distinguish between insufficient resources and inefficiency when evaluating school performances. To design effective policies and a better accountability system to improve student academic achievement, we must (1) understand what are the relevant and irrelevant inputs in the educational process (so that we can take into account differences in the relevant inputs when evaluating educational effectiveness); (2) understand how these inputs determine student achievement, especially whether lack of fi-

¹The other three districts are Jersey City, Newark and Paterson – source: <http://www.usatoday.com/story/news/nation/2013/03/25/camden-schools-takeover/2017259/>

nancial resources, if relevant, has detrimental effects on test scores (the effects would provide useful information to evaluate policy proposals for resource allocation) and (3) obtain accurate measures of inefficiency. These tasks – a study of (ir)relevant inputs and examination of the role of insufficient resources and inefficiency in the educational production process – are closely related and indeed are the goals of our paper.

Two related strands of the literature have examined these issues. Some researchers have examined the educational production function by examining various determinants of test scores such as class size and school expenditures. Other researchers have employed such methods as stochastic frontier analysis or its variants (based on the educational production function) to estimate and study school efficiencies (e.g., Collier and Millimet, 2009; Conroy and Arguea, 2008; Cooper and Cohn, 1997). Several (disturbing) findings typically emerge from these studies. First, a well-cited result from studies focusing on estimation of the determinants of test scores is that schooling inputs have limited success (e.g., Eherenberg and Brewer, 1994; Hanushek, 1996, 2003); such finding runs counter to the popular belief that making more resources available to schools improves student achievement. By contrast, the literature tends to find that environmental factors such as family backgrounds and neighborhood characteristics play a more important role in test score production (Cooper and Cohn, 1997). Even so, at most, modest associations exist between neighborhood characteristics and student achievement (Vigdor and Ludwig, 2010). Second, there exists inefficiency in the educational production process and schools are not performing to their potential (e.g., Conroy and Arguea, 2008; Cooper and Cohn, 1997).

In this paper, we employ stochastic frontier analysis to estimate school (in)efficiency. This approach is also called the educational production approach since it ties together estimation of an educational production function and efficiency in a unifying framework. Thus it is particularly suitable for the aforementioned goals of this paper. Moreover, this approach can adequately handle repeated measure (panel) data. It first entails measuring educational outputs and exploration of the underlying determinants/“inputs” of these outputs. It then measures the (in)efficiency as the distance between the maximum potential output (determined by the current educational technology and shocks), and the current observed output, given the amount of available inputs.

As a result, the key to estimation of inefficiency lies in estimation of the potential output and hence of the production process. Although the previous literature provides careful and important evidence on the effects of educational inputs, there are at least two gaps remaining. First, there may be heterogeneity in the returns to educational inputs. There is no reason to believe that the return to each input is constant across districts. Second, the existing educational production function literature relies mostly on parametric, linear regression

models (e.g., Conroy and Arguea, 2008). There has been evidence in the literature that the linearity assumption is not appropriate. For example, Figlio (1999) employs a more flexible translog production function and Eide and Showalter (1999) use a quantile regression approach. Both studies find supporting evidence for the existence of non-linearities. Although more flexible, these approaches are parametric and still impose potentially stringent assumptions. Furthermore, a parametric distribution is typically assumed on the errors and inefficiency in order to recover the efficiency measures. If the functional form assumptions do not hold, the parametric model will most likely lead to biased and inconsistent estimates of the return to each input and of efficiency scores. Biased estimates of efficiency scores could be misleading for policy-making. For example, we may unnecessarily penalize schools, teachers and/or students or call for costly reforms for perceived under-performing.²

In light of some of these potential shortcomings, we propose using a nonparametric kernel approach (see also Parmeter and Racine, 2012 or Martins-Filho and Yao, 2007). The approach relaxes the functional form assumptions associated with traditional parametric regression models and allows for potentially complicated interactions among the inputs of the model. Furthermore, the approach generates unique partial effects for each observation for each variable. This attribute enables us to make inference regarding heterogeneity in the returns. Unlike conventional nonparametric estimators, which typically assume that all variables are continuous in nature (e.g., Fan et al., 1996; Park et al., 1998, 2003), using discrete kernels, we incorporate discrete inputs (Racine and Li, 2004) into educational production functions and estimation of efficiency score.

Using data from public school districts in Illinois, our nonparametric analysis proceeds in three steps and yields three sets of results. These results highlight the three distinct features of public schools – irrelevant inputs, insufficient resources and inefficiency, as indicated in the title of this paper.

First, we employ automated bandwidth selection criteria to evaluate the relevance of variables in nonparametric settings (Hall, Li, and Racine, 2007). Our results show that *all*

²Efficiency scores have become widespread in regulated industries, and educational sector, especially in Europe. Benchmarking and efficiency analysis is routinely performed in water utilities, electricity and natural gas, rail transportation and postal services (Thanassoulis (2000a,b), Cubbin (2005), Dassler et al. (2006)). For example, on January 1, 2009, the German regulator (Federal Network Agency) initiated the Incentive Regulation Program which called to benchmark (in terms of efficiency) German electricity and gas network operators nationwide. Participation for the approximately 850 electricity and 730 gas companies was compulsory. To assess the efficiency score, the German regulator in this case employs both Stochastic Frontier Analysis (SFA) and Data Envelopment Analysis (DEA) methods and for each firm picks the one that reveals the largest efficiency level. The Department for Education and Skills (DfES) in U.K. and Economic Policy Committee of EU also have been interested in techniques such as Stochastic Frontier Analysis (SFA) to measure allocative and technical efficiency in secondary schools (Smith and Streets, 2006) and higher education (EU Economic Policy Committee Report, 2009), respectively.

of the commonly used “schooling inputs,” such as number of schools in the district, current per pupil expenditure and capital per pupil expenditure are irrelevant, while most of non-schooling inputs are relevant. This may not seem surprising given findings of insignificance for schooling inputs in the literature. However, we want to emphasize that our results are stronger because we do not impose functional form restrictions and allow for heterogeneity in returns in assessing the relevance of these variables. This point will be made more clear below.

Second, our estimates indicate that there exists substantial heterogeneity in the returns to relevant inputs across the sample. In particular, the effects of these inputs sometimes change signs. For example, we find that property tax caps have predominantly negative impacts (albeit sometimes insignificant) on test scores. This is the only relevant factor that is related to financial resources, and its prevailing negative association with test scores indicates that insufficient resources may be partially responsible for the lack of growth in test scores for many school districts. We further find that even though students tend to perform better in the districts with better educated populations, we find some districts where a better educated population does not improve test scores. We also find that the unemployment rate typically has a negative and significant impact on test scores, but the degree to which varies. These results help resolve the intriguing finding in the literature mentioned above – at most, modest associations exist between neighborhood characteristics and student achievement (Vigdor and Ludwig, 2010).

Finally, based on the estimated nonparametric production function, we obtain efficiency scores for each school and find that there exist inefficiencies among public schools. For this set of results, we have several findings. 1) We find that the measure of school effectiveness based only on test scores and our efficiency scores (after controlling for production technology and differences in resources) are related, but differ from each other substantially. Specifically, we find that there is a positive correlation between our efficiency scores and the level of test scores (implying that more efficient schools tend to produce higher test scores). However, such correlation is far from perfect (less than 0.4). This result suggests that after controlling for differences in inputs, schools with lower test scores are not necessarily performing less efficiently. 2) We generally find that the parametric specifications drastically underestimate the efficiency scores of schools. Parametric models could lead to potentially unnecessary and costly reforms intended to improve school efficiency, which may not lead to eventual gains in test scores. 3) Even though we find some evidence of negative impacts on the level of test scores, we are unable to find any significant difference in efficiency scores between schools which are financially constrained by property tax cap laws and those who are not. This finding may dispute some arguments which assume that schools can or will become more

efficient with these constraints.

2 Potential Endogeneity: Solutions and Limitations

In addition to functional form specification, a separate relevant empirical issue needs to be addressed. As with nearly all of the existing literature, estimation of production functions could be potentially biased if some unobservables are also correlated with observables. It is conceivable that some inputs are unobservable characteristics (such as ability and unmeasured policies) to econometricians and thus left out of the model. One may be tempted to think that the conventional instrumental variable (IV) approach may be a viable solution. This is *not* necessarily the case in our context for two reasons. First, any commonly used exclusion restrictions such as family background may themselves affect student academic achievement and thus cannot be valid IVs. As noted in a similar context by Dearden et al. (2002), “it is very hard to argue that any of the available background, family, or local variables determine school allocation but not educational attainment and wages. All such variables are potential inputs in the production of human capital.” Second, as noted in Henderson et al. (2011), the instrumental variable (IV) approach is not necessarily superior in the presence of heterogeneity. Since the conventional IV approach can only uncover the effect for the school districts whose inputs are affected by the IVs – the so-called local average treatment effects (Imbens and Angrist, 1994) – it can be argued that an IV is needed for every observation due to the substantial heterogeneity uncovered. This is unrealistic in practice. Moreover, even if such IVs are available, nonparametric estimation that could address both the endogeneity and heterogeneity issues simultaneously is, however, still far from practical implementation for micro-level studies like ours.

We follow the literature and respond to these potential shortcomings in two ways, although it is *a priori* unclear as to whether the magnitude (and hence empirical importance) of such potential bias indeed matters in the case of a nonparametric model. First, we include a broad set of control variables that are typically used to mitigate this issue. This is equivalent to advocating the assumption of “selection on observables” (as most studies do in this strand of the literature). Second, we focus on more aggregate data (school district level in our case), instead of individual level data, in only one state. Such practice is commonly adopted in the literature (e.g., Card and Rothstein, 2007 and references therein). As noted in Hanushek (2007), policies are usually set at the state level, and focusing on schools within the same state would allow us to compare schools that operate within the same basic policy environment” and thus to eliminate differences in unmeasured policies (e.g., Ruggiero, 1996). More important, the use of more aggregate data may also help to eliminate possible non-

random sorting in the distribution of student ability, a factor that is commonly perceived to bias the estimations. As noted in Card and Rothstein (2007), “students of differing abilities may sort to different schools or neighborhoods within a city, it is plausible that the distribution of potential abilities across metropolitan areas is closer to random (conditional on observed control variables)”. Evans et al. (1992) similarly use variation in average student characteristics to identify peer effects. In sum, we believe that these actions, although may not necessarily eliminate all possible biases, do address the type of the omitted-variable bias that is considered to be particularly important in this literature.

Finally, even considering potential limitations, we do hope that the main message of this paper – namely the existence of nonlinearities, heterogeneity and potentially underestimated efficiency scores resulting from stringent (linear) parametric assumptions – is clear and points to a promising venue to reconcile some puzzling results in the existing literature, but with important policy implications.

The remainder of the paper is organized as follows: Section 3 discusses both parametric and nonparametric estimation of an educational production function and determination of efficiency scores. Section 4 discusses the data and Section 5 presents the results. The final section discuss policy implications as well as possible future research.

3 Educational Production Functions and Technical Efficiency

In this section we describe the empirical methods used for estimating educational production functions and technical efficiency in a repeated measure (panel) data setting. There are generally two approaches for estimating efficiency: stochastic frontier analysis (SFA) and data envelopment analysis (DEA). Here we adopt SFA because of its ability to handle repeated measure data. As noted in Badunenko et al. (2012), there exists no “DEA estimator (with asymptotics proven) which can adequately handle [repeated measure] data”.

In SFA models, the production functions are first estimated and estimates of output oriented technical efficiency for each school district are then obtained as a by-product of the exercise. This basic framework assumes that we observe a cross-section of data for each of the N school districts for J repeated measures of (average) test scores. Quantities of d inputs are used to produce a scalar output through a production function. More specifically,

the production frontier model can be written as

$$\begin{aligned} y_{ij} &= y_{ij}^* - u_i \\ &= f(x_{ij}, \beta) + \varepsilon_{ij} - u_i \end{aligned} \tag{1}$$

where y_{ij} represents the level of output (typically average test scores) for school district i for the repeated measure j ($i = 1, 2, \dots, N, j = 1, 2, \dots, J$), y_{ij}^* represents the maximum potential output, given the amount of inputs (x_{ij}) available, and u_i is the non-negative technical inefficiency component of the error term (which is assumed to not vary over j). By construction, y_{ij}^* consists of two parts: (1) $f(\cdot)$ is the production function where β is a vector of unknown parameters; and (2) ε_{ij} represents the two sided noise component, a random technological shock to production.

Prior to continuing, it is worth noting that while test scores are *a* measure, they are not necessarily *the* measure, of schooling output. Test scores are the most commonly used measure of student performance by policy makers and academic researchers because they are most readily available for straightforward comparison across school districts. The use of it by no means reflects that performance in tests is the sole purpose of education.

3.1 Parametric Estimation

Although many methods exist for estimating a parametric production frontier model with panel data, here we choose one of the most popular methods and one that is comparable to our approach. Fixed Effects (FE) estimation of the production frontier, introduced by Schmidt and Sickles (1984), can be obtained, for example, from the log-linear Cobb-Douglas one-way error component model

$$y_{ij} = x_{ij}\beta + \alpha_i + \varepsilon_{ij}, \tag{2}$$

where $\alpha_i (= \alpha - u_i)$ is the district fixed effect. In other words, we assume that each district shares the same parametric technology in each time period, but that differences between them are captured by a location (district) effect α_i . Another popular parametric production function adopted in the literature is the translog production function (e.g., Figlio, 1999); this approach relaxes linearity and additivity assumptions in (2). Estimation of the coefficients (e.g., β in Equation 2) can be obtained, for example, by means of the within estimator.

Having obtained the estimated production function, district specific estimates of α are

then obtained by

$$\hat{\alpha}_i = \frac{1}{J} \sum_{j=1}^J (y_{ij} - \hat{y}_{ij}) \quad (3)$$

where \hat{y}_{ij} is the estimated output (e.g., $x_{ij}\hat{\beta}$ in the Cobb-Douglas case). Here we note that the estimates of the coefficients are consistent as $NJ \rightarrow \infty$, but $\hat{\alpha}_i$ is consistent only if $J \rightarrow \infty$. Further, as is common in this literature (Kumbhakar and Lovell, 2000), we estimate the district u_i by means of the normalization

$$\hat{u}_i = \hat{\alpha} - \hat{\alpha}_i,$$

where $\hat{\alpha} = \max_i \{\hat{\alpha}_i\}$ and if the left-hand-side variable is measured in logs, district specific estimates of technical efficiency are given by

$$\widehat{TE}_i = \exp(-\hat{u}_i).$$

3.2 Nonparametric estimation

As mentioned before, the (linear) parametric assumption in (2) may not be suitable for all panel data sets. In the education literature, the linear form of the production function has been statistically rejected (e.g., Figlio, 1999).³ As a result, we adopt a more general form for the production function, suggested by Kneip and Simar (1996):⁴

$$y_{ij} = g(x_{ij}) + \alpha_i + \varepsilon_{ij}, \quad (4)$$

where $g(\cdot)$ is an unknown smooth production function that each school district shares, but differences between them are captured by the location effect α_i .⁵ Identification of the indi-

³It is feasible to consider more flexible parametric models as well, but each of these will also require specification of the functional form.

⁴Note that our production function is very general, and alternatives such as semiparametric index models are only a special case of our model. For example, Projection Pursuit Regression (PPR) is one type of semiparametric models, which takes the form of additive index functions. While potentially more efficient, such models impose additive separability and hence restrict generality. As we will show below, incorrect restrictions on functional forms could lead to severely biased estimates of both partial effects and efficiency. Moreover, estimation of gradients (partial effects) for the PPR is generally not as straightforward, and when there is more than one additive functions, the PPR model generally lacks interpretability. Our model is therefore more preferred.

⁵Equation (4), as it is written, is a special case of the partially linear model attributed to Robinson (1988) where the α_i can be treated with dummy variables. It could be argued that this model is semiparametric as one component is nonparametric ($g(\cdot)$) and the other one is parametric (α_i). An empirical example of a repeated measure problem estimated via the estimator of Robinson (1988) can be found in Millimet, List and Stengos (2003). Our estimation method of $g(\cdot)$ will essentially ignore the additive structure by placing a mean zero assumption on the α_i , similar to that in Kneip and Simar (1996).

vidual effect requires the “average” of α_i to be equal to 0.

The estimator employed here differs from Kneip and Simar (1996) in two important ways. First, we consider an estimator of the production function which allows for discrete inputs. Conventional nonparametric analysis (e.g., Park and Simar, 1994; Sickles, 2005; Kneip and Simar, 1996) assumes that all inputs are continuous in nature. It is common in educational research for inputs to be discrete, for example, one of the schooling inputs considered in our application is the number of schools in the district.⁶ To address this issue, we use discrete kernels to incorporate discrete inputs correctly (e.g., Racine and Li, 2004).

Second, Kneip and Simar (1996) use local-constant least-square (LCLS) to estimate the production function, but here we suggest using the local-linear least-squares (LLLS) estimator. Not only does the local-linear estimator give more reliable estimates of $g(\cdot)$, it also allows for estimation of both the production function and its gradient in one step. The gradient measures the marginal effect of the input and allows us to immediately examine potential heterogeneity in returns to each input.

3.2.1 Nonparametric Estimation with Mixed Data Types

Here we outline the kernel estimator. The estimator employed here is essentially a pooled version of the LLLS estimator in Racine and Li (2004). We allow the vector x_{ij} to be composed of three different types of regressors: continuous, unordered discrete and ordered discrete variables. In particular, $x_{ij} \equiv [x_{ij}^c, x_{ij}^u, x_{ij}^o]$, where x_{ij}^c is a vector of continuous regressors, x_{ij}^u is a vector of regressors that assume unordered discrete values, and x_{ij}^o is a vector of regressors that assume ordered discrete values. We define d_c as the number of continuous inputs. The number of unordered and ordered discrete inputs are defined as d_u and d_o , respectively. Note that the total number of inputs $d = d_c + d_u + d_o$.

Taking a first-order Taylor expansion of (4) with respect to x^c yields

$$y_{ij} \approx g(x) + (x_{ij}^c - x^c)\beta(x) + \alpha_i + \varepsilon_{ij},$$

where $\beta(x)$ is defined as the partial derivative of $g(x)$ with respect to x^c . Note that the Taylor expansion can only be taken with respect to the continuous inputs.

⁶Many papers in the (nonparametric) literature have either used partially linear models, treated the discrete inputs as continuous, ran separate regressions for each value of the discrete regressor or excluded them altogether. Sometimes this is harmless. However, sometimes the nature of the categories are too important to ignore and could indeed affect the results.

The LLLS estimator of $\delta(x) \equiv [g(x) \beta'(x)]'$ is given by

$$\begin{aligned}\widehat{\delta}(x) &= [\widehat{g}(x), \widehat{\beta}'(x)]' \\ &= (X'K(x)X)^{-1} X'K(x)y,\end{aligned}\tag{5}$$

where X is a $NJ \times (d_c + 1)$ matrix with ij th row being $(1, (x_{ij}^c - x^c))$ and $K(x)$ is a $NJ \times NJ$ diagonal matrix of kernel weighting functions for mixed continuous and discrete data with bandwidth vector $(h, \lambda^u, \lambda^o) = (h_1, \dots, h_d, \lambda_1^u, \dots, \lambda_{d_u}^u, \lambda_1^o, \dots, \lambda_{d_o}^o)$ for the continuous and discrete variables, respectively. We use a product kernel function specifically designed for each type of input: the second-order Gaussian kernel for continuous inputs, and the kernels for unordered and ordered discrete inputs discussed in Aitchison and Aitken (1976).

The local-constant estimator (Li and Racine, 2004), which will be used to remove irrelevant inputs via the cross-validation procedure, can be obtained similarly via replacing the data matrix X with a column of ones (in other words, no first order Taylor expansion).

3.2.2 Efficiency scores

Following the practice above, having determined an estimator for the production function ($\widehat{g}(\cdot)$), estimators of α for each i can be similarly obtained as

$$\widehat{\alpha}_i = \frac{1}{J} \sum_{j=1}^J [y_{ij} - \widehat{g}(x_{ij})].\tag{6}$$

Again, we estimate the district u_i by means of the normalization

$$\widehat{u}_i = \widehat{\alpha} - \widehat{\alpha}_i,$$

where $\widehat{\alpha} = \max_i \{\widehat{\alpha}_i\}$. District specific estimates of technical efficiency are given by

$$\widehat{TE}_i = \exp(-\widehat{u}_i).$$

Comparing Equations (3) and (6), we can see the difference in efficiency scores stem from the difference between $x_{ij}\widehat{\beta}$ (the predicted values from parametric regressions) and $\widehat{g}(\cdot)$ (the predicted values from nonparametric regressions). In other words, the (linear) parametric approach treats any nonlinear component that the regression fails to capture as part of the inefficiency. One may be tempted to think that taking into account the heterogeneity and nonlinearity would decrease the inefficiency scores. This is not necessarily true, especially when we have more than one input. The changes in the efficiency scores depend on whether

or not the nonlinear terms for each input will cancel one another out. The direction of such changes is an empirical question to be examined below.

3.2.3 Bandwidths and relevance of variables

Prior to continuing, another practical detail is worth emphasizing. Estimation of the bandwidths can be obtained by using least-square cross-validation (LSCV). In short, the procedure chooses bandwidths $(h, \lambda^u, \lambda^o)$ such that they minimize the LSCV function

$$CV(h, \lambda^u, \lambda^o) = \sum_{j=1}^J \sum_{i=1}^N [y_{ij} - \hat{g}_{-i}(x_{ij})]^2,$$

where $\hat{g}_{-i}(x_{ij})$ is computed by leaving out, as in Kneip and Simar (1996), all the J observations (y_{ij}, x_{ij}) of the i th district (leave-one district-out estimator).

Hall, Li and Racine (2007) show that a useful feature of the (leave-one-out) LSCV procedure is its ability to *statistically* determine the relevance of the variables of interest *simultaneously*, as opposed to sequential examination.

For continuous regressors, in the local-constant case, a bandwidth equal to the upper bound implies that the variable is irrelevant. In the local-linear case, a bandwidth equal to the upper bound determines that the (continuous) variable enters in linearly. The upper bound for the bandwidth on a continuous regressor in either case is infinity. This is impossible to observe in practice. However, when using a Gaussian kernel function, any bandwidth in excess of a few standard deviations of the regressor gives essentially equal weight to all observations. In other words, in the local-constant setting, the local average with respect to that variable is actually a global average of the left-hand-side variable and hence the regressor (essentially) has no impact on the conditional mean. In the local-linear setting, all observations are given equal weight and hence the regressor enters the model (essentially) in a linear fashion. Thus, we follow the suggestion of Hall, Li and Racine (2007) and use two standard deviations of the regressor as the bound for relevance/linearity. Thus, if any bandwidth on a continuous regressor exceeds two standard deviations of its associated variable, we conclude that it enters in an irrelevant fashion (in the local-constant setting) or linearly (in the local-linear setting).

For discrete variables, the bandwidths, either for local-constant or local-linear estimation, indicate which variables are relevant, as well as the extent of smoothing in the estimation. From the definitions for the ordered and unordered kernels (see appendices), it follows that if the bandwidth for a particular unordered or ordered discrete variable equals zero, then the kernel reduces to an indicator function and no weight is given to observations for which

$x_{ij}^o \neq x^o$ or $x_{ij}^u \neq x^u$. On the other hand, if the bandwidth for a particular unordered or ordered discrete variable reaches its upper bound, then equal weight is given to observations with $x_{ij}^o = x^o$ and $x_{ij}^o \neq x^o$. In this case, the variable is completely smoothed out (and thus does not impact the estimation results). For both unordered discrete variables, the upper bound is $(c - 1) / c$, where c is the number of distinct values the discrete regressor takes. For ordered discrete variables, the upper bound is unity. See Hall, Li and Racine (2007), Henderson and Parmeter (2015) and Li and Racine (2006) for further details.

3.2.4 Proposed Implementation

In our application, we will first use the local-constant estimator with LSCV bandwidths to determine relevance of our inputs. We will then remove irrelevant inputs, re-estimate the LSCV bandwidths for the relevant variables only and use a local-linear regression to examine the partial effects as well as determine our efficiency scores.⁷

4 Data

Our data include public school districts in the state of Illinois and which are obtained from Millimet and Collier (2008). As noted in Hanushek (2007), policies are usually set at the state level, and focusing on schools within a state would allow us to compare schools that “operate within the same basic policy environment” and thus to eliminate differences in unmeasured policies (e.g., Ruggiero, 1996). The use of more aggregate data (such as ours at the school district level) is commonplace and can be advantageous in this context (e.g., Card and Rothstein, 2007 and references therein) because such practice may help to eliminate possible non-random sorting, a common challenge facing all empirical analysis. For example, Card and Rothstein (2007) argue that “students of differing abilities may sort to different schools or neighborhoods within a city. It is plausible that the distribution of potential abilities across metropolitan areas is closer to random (conditional on observed control variables)”.

Summary statistics for the variables discussed below can be found in the Appendix (Table 8). The outcome of interest is test scores in reading and mathematics administered by the Illinois State Board of Education annually in grades three, six, eight and ten; such tests

⁷Note that removal of irrelevant variables is automated in local constant estimation. Therefore, as clearly shown in Hall et al. (2007), pretesting is unnecessary when one employs cross-validated bandwidth selection in nonparametric estimations. As such, the first-stage estimation in our case should not be considered as a pre-testing stage, and our least-squares estimation should also be considered as a refined estimation procedure that provides direct estimation of gradients. Further, we employ a wild bootstrap procedure to obtain standard errors; this procedure is preferred since it takes into account variations in any multi-stage estimations and is robust to heteroskedasticity that would otherwise render a standard residual bootstrap invalid for statistical inference (Cameron and Trevedi, 2005).

are “intended to provide a common yardstick for monitoring school and district progress”. The tests are scored on a scale of zero to 500. These eight test scores are available for both the 1996-1997 and 1997-1998 academic years.⁸ Following Millimet and Collier (2008), the districts with missing information in either year are excluded, and the resulting data are a balanced-panel sample with $N = 850$ school districts each with $J = 16$ test scores. Following Millimet and Collier (2008), we pool together the scores of different subject test in estimations. However, this should not be considered restrictive. Recall that our nonparametric estimation allows for arbitrary interactions among *all* variables. Thus, all the relevant inputs could in effect have differential effects on test scores for each subject. Moreover, pooling together all the subject tests increases the number of repeated measures and may improve our estimates of efficiency in practice.

As noted in Figlio (1999), “student learning is derived from many sources – including families and peers – and not just schools.” The inputs of the production function (x_{ij}) typically include schooling inputs (which could be controlled by administrators) and non-schooling inputs (which are beyond the control of district administrators) such as family background, peer group characteristics, and local environmental factors (Vignoles et al., 2000). As Ruggiero (1996) argues, “it is essential to control for environmental factors that affect the translation of school inputs into student achievement”.

In our analysis, we utilize the following inputs at the district level in the estimation of the educational production function, and these inputs could be loosely categorized into two groups: schooling v.s. non-schooling inputs. Schooling inputs include (1) Resources (current per pupil expenditure, capital per pupil expenditure, full-time equivalent (FTE) teachers) (2) District Organization (the number of schools, Local Education Agency (LEA) administrators). These inputs are directly manipulable by policymakers. The number of FTE teachers, schools, and LEA administrators are discrete variables in nature.

Non-schooling inputs include district-level characteristics and demographic information (total population, median household income, number of persons aged 20+ without a high school diploma, number of persons aged 20+ with a college education, unemployment rate, percentages of children aged 4-19 whose primary language is something other than English, percentages of occupied housing, and percentages of owner-occupied housing). Employing a variety of statistical techniques, Gyimah-Brempong and Gyapong (1991) argue that background and demographic variables are significant factors in the production of education, and that exclusion of these factors could result in a misspecified production function. More importantly, there could be important inputs that are unobservable and left out of the model,

⁸Here we choose data which was previously used in the literature in order to facilitate direct comparisons to the existing results. Specifically, for our goal to highlight functional form and heterogeneity.

and thus it is important to condition on a broader set of variables. We indeed find that these are relevant inputs. Deller and Rudnicki (1993) similarly find that family backgrounds – proxied by adult education levels and per capita income of the communities composing the school district – influence test score production. These authors further emphasize the role of peer influences, capturing student aspirations and ambitions. They argue that the level of the unemployment rate in the area in which the school is located is highly related to student aspirations and is thus a good proxy for peer influences. Moreover, unemployment may also reflect time constraints facing parents and hence potential parent effort. Parental involvement has been shown to be an important input in the educational production function (e.g., Houtenville and Conway, 2008).

In addition to the aforementioned factors, the literature has also pointed out the importance of institutions in the production of education. Recall that we focus only on school districts in Illinois, and as a result, the difference in these institutional variables should be relatively small. However, in our analysis, we do include an institutional variable at the local level that could potentially affect the production of test scores. In particular, we include fiscal constraints - whether the Property Tax Extension Limitation Law (PTELL), commonly called “tax caps”, is implemented. The law places a cap on the *growth* of property tax revenue, which regulates that the growth of tax revenue cannot exceed the growth of the inflation rate. Unlike other levels of local government, school districts do not have any other sources of revenue, and hence they are more dependent on property tax revenue. Schools under PTELL counties are thus more financially constrained. Millimet and Rangaprasad (2007) and Millimet and Collier (2008) show that this financial constraint could play an important role in analysis of an educational production function and efficiency. This variable could also proxy for differences in resources other than the ones discussed above. The measure used is a binary variable equal to 1 if PTELL is implemented and zero otherwise.

In addition, we also include test ID to indicate the type of test scores (an unordered discrete variable taking values 1-8 corresponding to a particular subject and grade level) and a variable which indicates the year the test was taken (an ordered discrete variable).

5 Estimation results

We first examine the bandwidths from each cross-validation procedure. Second, we will focus on the partial effects of the educational production function. We then turn our attention to the efficiency scores.

5.1 Bandwidths and relevance of variables

Table 1 gives the bandwidth for each of the regressors in our study. First, we find that nearly all non-schooling inputs (e.g., median household income and unemployment rate) are relevant determinants of students’ academic achievement. This result is not surprising. Non-schooling inputs such as families and peer effects have been found to be important since the publication of the Coleman (1966) report (Harris, 2010), and the importance of these inputs, especially family background, is evident in more recent studies (Rothstein, 2004).

Our results with regard to the (percentage) educated population and income are consistent with prior findings that socioeconomic status is associated with educational outcome (e.g., Vigdor and Ludwig, 2010). However, these results seem suggestive that among the components of socioeconomic status, education may matter more than income. Specifically, we find that controlling for education, median household income in the school district is irrelevant, while education, measured by the number of persons with a bachelors degree (noting that we also control for population), matters for student achievement. This result is consistent with Gyimah-Brempong and Gyapong (1991). The authors similarly find that while family backgrounds play an important role in the production process, “inclusion of adult educational levels in the equation is sufficient” (Cooper and Cohn, 1997). In their paper, adult education is defined as the percentage of adults with 12 or more years of formal schooling in the district whereas our measure is the number of persons with a bachelors degree. We also find PTELL to be relevant. This policy variable is highly debated and its impact on test scores will be examined in more detail later.

In stark contrast to non-schooling inputs, *none* of the traditional schooling inputs are relevant. While these results are (qualitatively) in agreement with the literature, our results are much stronger. Recall that in parametric models, relevance of a variable is usually tested by statistical significance of the single coefficient on that variable (obtained, say via OLS). This coefficient represents the average effect of this variable. In the presence of heterogeneous effects, where both negative and positive effects are possible across districts, it is possible that these opposing effects could offset each other and then the average effect is zero. This could result in a statistically insignificant coefficient and in turn lead one to falsely conclude that the variable is irrelevant. However, this is not the case here since we do not impose any functional form restrictions and allow for heterogeneity in returns in assessing the relevance of these variables. Thus, it is a rather strong result that *all* schooling inputs are found to be irrelevant, even in such a general setting.

5.2 Partial effects of Relevant Inputs

Table 2 presents the results for the gradient estimates from the local-linear regression. Non-parametric estimation provides observation-specific partial effects and standard errors. To conserve space, we present the nonparametric estimates corresponding to the 10th, 25th, 50th, 75th and 90th percentiles of the estimated gradient distributions, along with each corresponding wild bootstrapped standard error in parentheses. The results for the rest of the distribution as well as the parametric results are available from the authors upon request.

Table 2 first reports the bandwidths for each relevant input as well as their corresponding upper bounds. As we can see, nearly all bandwidths (except the bandwidth for percentage of occupied housing in the district) are smaller than their upper bounds. As discussed above, this suggests that these inputs indeed enter the production function nonlinearly, and hence linear, separable parametric models are likely misspecified.⁹

Table 2 reveals an important message: precision set aside, there exists a large extent of heterogeneity in partial effects for each input. Not only do these partial effects vary in magnitudes, but also in signs. This heterogeneity likely comes from either nonlinearities and/or interactions between variables. We note that the types of nonlinearities and interactions are generally unknown which makes this difficult to model in a parametric setting. As we shall see below, even a relatively flexible translog production function cannot capture these nonlinearities and interactions sufficiently and leads to downward biased estimates of efficiencies.

Several findings are noteworthy. First, we find that for the majority of school districts, financial constraints (PTELL) typically exert a negative impact on students' educational outcomes. Although some are insignificant, the majority of school districts have a negative coefficient on PTELL. This result is consistent with the growing evidence that imposition of tax limits results in reductions in student achievement (e.g., Downes and Figlio, 1999). This variable captures the financial constraint facing a school district. This negative association indicates insufficient resources could be partially responsible for the lack of growth in test scores. One additional possibility is that tax limits could be generally imposed in districts that place less value on public education. This result also suggests there may be room for policy to improve students' academic achievement. There could be two reasons why tax

⁹The local-linear estimator in equation (5) is a special case of weighted least-squares, where the weights are determined by the distance between the point of interest and the data point (where the relative distance is determined by the bandwidth). When the bandwidth gets large, all data points are given the same weight and we are back to the standard OLS estimator (i.e., linearly separable). More formally, as $h \rightarrow \infty$, $K(\cdot) \rightarrow K(0)I$ (a constant times an identity matrix) and cancels out. Equation (5) essentially becomes $\hat{\delta}(x) = (X'K(x)X)^{-1}X'K(x)y = (X'K(0)IX)^{-1}X'K(0)Iy = (X'X)^{-1}X'y$, where the latter is the familiar OLS estimator.

limits have such negative effects on test scores. First, it could be that financial resources matter through channels other than the traditional inputs that we examined above. Second, even though the literature generally finds that the levels of school inputs do not affect student performance, as evident above, Figlio and Rueben (2001) have shown that tax limits are associated with reduced quality of these inputs, which could in turn reduce student performance. However, unlike prior studies, we also recover a proportion of districts where these constraints play a positive role in student’s academic achievement. This result could be masked by simple OLS estimates which are an average of these positive and negative effects. In light of the large literature on strategic competition among schools (e.g., Millimet and Collier, 2008), such positive effects may reflect competitive pressure imposed by districts without limits. Fiscally constrained schools may be forced to be more efficient by reallocating resources productively. We will examine this matter later. The magnitudes of the effects also vary across districts. For example, the estimated effect at the 10th percentile (-0.047) is roughly 2 times larger than at the 25th percentile (-0.024). Note that these estimates are elasticities. To put these results in perspective, consider that the average district test score is 270.168 with a standard deviation of 32.276 (Table 8). Thus, imposing tax limits in a school district with average test scores would reduce test scores by less than one point ($= (270.168 \times 0.047\%)$). This is not a large effect and could be due to the opposing effects of the tax limits on test scores discussed above.

Second, students tend to perform better in districts with better educated populations. Specifically, we find that the effect of the percentage of population with bachelors degrees on students’ outcomes are generally positive. However, we do find a small fraction of the districts where a better educated population does not necessarily improve test scores, albeit insignificantly. This is consistent with theoretical ambiguities regarding the effects of having neighbors of higher socioeconomic status. The positive effects may reflect “own effects” (e.g., the effect of the student’s parental education in the educational production function) through mechanisms such as parental involvement (e.g., Houtenville and Conway, 2008). It could also be beneficial because of mechanisms such as positive role models and reduced exposure to violence (Kling et al., 2007). On the other hand, it could be adverse because children may face competition with advantaged peers and discrimination. Moreover, even for those districts with positive effects, the magnitudes of the effects again vary. For example, the coefficient at the (unreported) 95th percentile of the distribution of the partial effects is at least twice as large as the coefficient at the 75th percentile (0.062 at the 95th vs. 0.029 at the 75th).

Third, we also find that the percentage of kids whose primary home language is not English has a large positive effect on educational outcomes in a majority of school districts.

This result is broadly consistent with the existing literature, which fails to find harmful effects of immigration on natives' educational outcomes (e.g., Liu, 2000), and even finds slightly positive effects (e.g., Neymotin, 2009). Chin et al. (2012) similarly find that the number of limited English proficient (LEP) children does not affect the LEP students but has significant positive spillover effects on non-LEP students. Such a result could be explained by the fact that resources provided for LEP students are also beneficial to non-LEP students. Nevertheless, we find that the percentage of children whose home language is not English does harm student achievement in some districts (marginal significance). Such a result is masked by ordinary regressions that examine only average effects.

Finally, we find that for a majority of school districts, the unemployment rate predominantly has a negative effect on student achievement, while housing generally has a positive effect. Further, the gradients for the unemployment rate is larger in magnitude for lower percentiles of the distribution of the gradients. These results may be explained by the fact that high mobility due to lack of affordable housing and unemployment could depress achievement for both students who move and students who do not move (Rothstein, 2004, p.46). Our results suggest that housing-mobility interventions such as the Moving-To-Opportunity experiment that generate changes in neighborhood characteristics could be quite effective in improving student outcomes for some districts, although not for all (e.g., Kling et al. 2007).

Note that as mentioned above, our approach allows for arbitrary interactions among all variables and permits differential effects of relevant inputs on test scores for different subjects. We formally test this by performing pair-wise Kolmogorov-Smirnov tests of equality of the distributions of partial effects between different subjects for each relevant input variable. The results are presented in Tables (3) and (4). We easily reject the equality for nearly all pairs. These results confirm our conjecture.

In sum, we find that there exists substantial heterogeneity in the effects of various factors on test scores. The findings of large heterogeneity in impacts helps to reconcile the fact that the literature consistently find at most modest associations between neighborhood characteristics and student achievement (Vigdor and Ludwig, 2010). These findings also highlight the importance of employing nonparametric techniques to complement the existing approaches to investigate the factors contributing to student achievement.

5.3 Technical Efficiencies

We now turn to the technical efficiency scores. We compare our measures to those obtained using Cobb-Douglas and translog specifications. The summary statistics of the efficiency scores for each approach are given in Table 5. We present the 25th, 50th, 75th, percentiles of

the estimated efficiency scores as well as the minimum, mean, and maximum values.

The nonparametric estimates of efficiencies are larger than their parametric counterparts. The magnitudes of the differences are also sizable. For example, the minimum efficiency score for the nonparametric approach is 0.265, while the minimum scores are 0.038 and 0.003 for the Cobb-Douglas and translog functions, respectively. The mean efficiency score for the nonparametric approach is twice as large as either parametric result. Further, the spread of the efficiency scores are much smaller for the nonparametric approach than for the parametric ones. Specifically, the coefficient of variation for efficiency scores using the nonparametric approach is only 5.100, while that for the Cobb-Douglas and translog functions are 38.617 and 44.719, respectively. Note that the coefficient of variation for the level of test scores is 16.300.

To further examine the relationship between different approaches to measure efficiency (the observed level of tests vs the parametric SFA vs the nonparametric SFA), we examine the correlation rankings in Table (6). Here we consider three different correlation coefficients: Kendall, Pearson and Spearman. While there is little correlation between efficiency scores obtained from the parametric approaches and the observed test scores, there is a positive correlation between our nonparametric efficiency scores and the observed test scores. Our result suggests that more efficient schools tend to produce higher test scores. However, such correlation is far from perfect (the estimated correlations are $0.2 < \rho < 0.4$). This result implies that the observed test scores, without taking into account the differences in inputs, are not necessarily an accurate measure of school efficiency. We should not simply use the observed test scores as a ranking of school efficiency. Part of the differences in test scores is also attributable to the differences in the inputs between school districts (which may be beyond the control of policymakers). The finding that there is a positive, but weak relationship between efficiency and test scores may not surprise many academics. However, as mentioned in the introduction, many policy makers do consider lack of growth in observed test scores as a measure of ineffectiveness, a view that we strongly reject.

We notice that the parametric approaches produce rather similar efficiency scores. All three measures consistently show a very high correlation between the rank estimates of the parametric approaches. By contrast, the correlations between our nonparametric approach and parametric scores are relatively small. These results are even more evident when examining the densities of technical efficiencies (Figure 1). We find that the distribution of technical efficiencies from the parametric approaches are surprisingly close to one another, while that from the nonparametric approach is very distant from these parametric results. Stronger visual evidence is seen in Figure 2 where we look at the empirical cumulative distribution functions. These are complemented by Kolmogorov-Smirnov tests (Table 7) where

we reject equality between each of the parametric and nonparametric distribution estimates of technical efficiency and at the same time fail to reject first-order uniform dominance of the nonparametric efficiency estimates over each of the corresponding parametric estimates.¹⁰

Given the findings in Gong and Sickles (1992), this suggests that the parametric estimates are downward biased. This is not only important from an econometric standpoint, but it may also be important for policy. If we believe that schools are under-performing, we may unnecessarily penalize schools, teachers and students.

5.4 PTELL as an Illustrative Example

Here we further highlight how parametric measures of inefficiency in public school districts may be problematic and provide misleading messages. Also, we briefly show how researchers and policy makers can benefit from our non-parametric efficiency measurements by combining them with further regression analysis. Specifically, we focus on a particularly important factor in determining school performance – property tax caps. Unlike other relevant variables such as educational levels in a district, this is one of the relevant inputs that policymakers can more easily manipulate. Examination of the *association* between property tax caps and efficiency could thus have relevant policy implications.

The proponents of PTELL argue that financially constrained (PTELL) schools will be forced to run schools more efficiently. Here, we examine this claim by looking at the efficiency scores by PTELL status. In Figure 3, we separate distributions of efficiency scores based on whether PTELL was equal to zero or unity. The distributions of efficiency scores (in the nonparametric case) are essentially identical. Our nonparametric results are in stark contrast to the parametric results obtained from Cobb-Douglas (Figure 4) and translog (Figure 5) models.

Our nonparametric result suggests that the fiscal constraints do not make these schools more efficient. In other words, the negative shift in the production frontier brought about by property tax caps (suggested by the negative effects of PTELL on the level of test scores) is not met with a movement toward the frontier. In fact, a simple check of correlation shows very little or even a negative relationship (albeit insignificant) relationship between PTELL and efficiency scores. This result is in stark contrast to the beliefs held by some researchers (Brennan and Buchanan, 1979) and many voters (e.g., Citrin, 1979) that tax and expenditure limits would make school districts eliminate waste. Recall that the majority

¹⁰Note that our inferential procedures for Kolmogorov-Smirnov tests do not take into the fact that efficiency scores are estimated. More complicated bootstrapping procedures involving estimation of all nonparametric results could be done but are computationally challenging (see, e.g., Henderson and Maasoumi, 2014). However, note that regardless of the inferential procedures used, the empirical observation of the existence of the first-order dominance should (in our case) will remain unchanged.

(about 63%) of the schools constrained by PTELL experience reductions in test scores to varying degrees. This indicates that the lack of growth in tax revenues leads to less resources devoted to education, thereby negatively affecting students' school performance. This result also suggests that financial resources may still matter for student achievement to some extent in the inputs that are not necessarily captured by traditional schooling inputs. Such a result implies that there may be some room for policy makers to improve students' test scores.

In our analysis, if we run a simple regression of whether a school has a positive coefficient on PTELL on technical efficiency scores, the coefficient is positive and statistically significant, suggesting that efficient use of resources could potentially improve student achievement.¹¹ The question is: why would administrators in these districts *not* have an incentive to be more efficient to offset the negative impact? One possible explanation is put forth in the literature (Downes et al. 1998): budget-maximizing administrators could use the decline in student performances to argue for more resources and thus do not have incentives to perform more efficiently. Consistent with this explanation, Figlio and O'Sullivan (2001) provide evidence that cities subject to tax limits strategically manipulate the provisions of inputs (particularly by cutting largely instruction-related inputs, but not administrative inputs) to encourage local voters to override the limit.

As emphasized above, our current analysis of the relationship between PTELL and efficiency is only illustrative and can only be interpreted as *correlation*, instead of *causality*. In the future, a more systematic and rigorous analysis of this question is warranted, but it should also take into account the possibility that misspecification of the first-stage estimation of efficiency scores could drastically change the results. It is our hope that this point is made clear.

6 Policy Implications and Future Research

Having obtained all these results, it is worthwhile to step back to review some of the important findings and discuss their policy implications more coherently. Along the way, we would also like to point out possible future research. As we have emphasized, our results suggest that public school districts in Illinois are characterized by the three *Is*: irrelevant inputs, insufficient resources and inefficiency. Even though our study focuses only on the state of Illinois, by our review of the current literature, we suspect that these are characteristics of many public school districts in the U.S. While these results may seem to paint a bleak picture of the current state of the public education system, in our view, they also provide some

¹¹Note that the criticism of two-stage estimators discussed in Wang and Schmidt (2002) is not relevant here as we are not in the maximum likelihood framework.

opportunities and possibilities for us to improve student academic achievement. Specifically, according to our results, we believe that there are two types of policies that should be carefully explored and warrant further scrutiny: one is to improve the relevant inputs in the production of test scores, and the other is to reform school districts to utilize resources more efficiently by holding them accountable.

Recall that we find that most traditional schooling factors are irrelevant (inputs) in the production of test scores. This is an established result in the literature. Thus, it may seem unsurprising to many readers familiar with the field. However, as we mentioned, the fact that it is irrelevant even after considering all possible nonlinear effects and interactions with other inputs (allowed in our nonparametric setting) is a rather strong and indeed somewhat surprising result. Also, the literature often cites the possible explanation that failure to find any significant effects may be a result of ignoring the possible interactive effect between the level and quality of these inputs in the estimations. Our nonparametric approach allowing for this type of interactions still fails to find any relevance, which could suggest that improving the quality of these traditional schooling inputs may not necessarily increase test scores. Thus, for the first type of policies, this result alone helps to at least narrow down the areas where policy may be effective and suggests that the focus of policy should shift away from traditional schooling inputs to non-traditional schooling inputs or non-schooling inputs.

The literature often concludes that money does not matter because traditional schooling inputs, which could be purchased by “money”, are irrelevant in the educational production. However, our finding that financially constrained school districts perform worse indicates that money does matter (at least to some extent). These two results together suggest that we should allocate more (financial) resources to school districts to improve non-traditional schooling inputs (those that are not typically included in publicly available data and thus not captured in our analysis). Specifically, future research should focus on identifying those relevant non-traditional schooling inputs that could be potentially affected by PTELL. For example, Parker (2009) finds that PTELL is associated with reduced extracurricular activity offerings, which have been shown to be associated with student academic achievement (e.g., Everson and Millsap, 2005; Holloway, 2000). Our results also highlight the importance of neighborhood characteristics in educational production. Such results point to possible non-schooling inputs that we should take into account in designing policies. For example, the results seem to lend support to residential mobility programs such as the *Gautreaux* program and the Moving to Opportunity program that are aimed to relocate poor families to “better” communities via housing vouchers. Regardless of policy directions, our finding of substantial heterogeneity in partial effects suggest that there are no one-size-fits-all type of policies. Future research should focus on possible heterogeneity in policy impacts and study

the sources of heterogeneity. Policies should also be accordingly designed to target students or families that may actually benefit.

Given the uncovered inefficiency among school districts, it is indeed important to hold them accountable for making more efficient utilization of resources. However, we should be cautious about how we measure school efficiency. Our results stress that in designing any accountability system, the final outcomes (such as test scores) may not be the best measure, as they do not take into account differences in relevant inputs. We believe that SFA is a useful tool, but we should always be careful about the functional assumptions in practice. Given the drastic differences between our nonparametric approach and parametric ones, it is our hope that some of the important studies of school efficiencies will be re-visited using our approach to provide information required for sound policies; and that our nonparametric approach may be of use when devising measures of school efficiency in policies.

Finally, we document a large heterogeneity in how different factors affect student achievement. A more comprehensive understanding of what drives such heterogeneity is warranted and could be a fruitful avenue for future research.

References

- [1] Aitchison, J. and C.G.G. Aitken. (1976). “Binary Discrimination by Kernel Method,” *Biometrika*, 63, 413-420.
- [2] Badunenko, O., D.J. Henderson and S.C. Kumbhakar (2012). “When, where and how to perform efficiency estimation,” *Journal of the Royal Statistical Society, Series A*, 175(4), 863-892.
- [3] Brennan, G. and J. Buchanan. (1979). “The Logic of Tax Limits: Alternative Constitutional Constraints on the Power to Tax, ” *National Tax Journal Supplement*, 32(2), 11C22.
- [4] Cameron, C. A. and Trivedi, P. K. (2005). *Microeconometrics: Methods and Applications*, Cambridge University Press, New York, NY.
- [5] Card, D. and J. Rothstein (2007). “Racial Segregation and the Black-White Test Score Gap,” *Journal of Public Economics*, 91(11-12), 2158-2184.
- [6] Chin, A. N.M. Daysal and S.A. Imberman (2012). “Impact of Bilingual Education Programs on Limited English Proficient Students and Their Peers: Regression Discontinuity Evidence from Texas, ” *NBER Working Papers # 18197*.
- [7] Citrin, J. (1979). “Do People Want Something for Nothing: Public Opinion on Taxes and Government Spending, ” *National Tax Journal Supplement*, 32(2), 113C29.
- [8] Cubbin, J. (2005). “Efficiency in the water industry, ” *Utilities Policy* 13(4), 289C293.
- [9] Collier, T. and D.L. Millimet (2009). “Institutional Arrangements in Educational Systems and Student Achievement: A Cross-National Analysis,” *Empirical Economics*, 37, 329-381.
- [10] Conroy, S.J. and N.M. Arguea (2008). “An Estimation of Technical Efficiency for Florida Public Elementary Schools,” *Economics of Education Review*, 27, 655-663.
- [11] Cooper, S.T. and E. Cohn (1997). “Estimation of a Frontier Production Function for the South Carolina Educational Process,” *Economics of Education Review*, 16, 313-327.
- [12] Dassler, T., Parker, D. and D.S. Saal (2006), “Methods and trends of performance benchmarking in UK utility regulation, ” *Utilities Policy* 14(3), 166C174.

- [13] Duncombe, W., J. Miner and J. Ruggiero (1997). "Empirical Evaluation of Bureaucratic Models of Inefficiency," *Public Choice*, 93, 1-18.
- [14] Deller, S.C. and E. Rudnicki (1993). "Production Efficiency in Elementary Education: The Case of Maine Public Schools," *Economics of Education Review*, 12, 45-57.
- [15] Dearden, L., J. Ferri and C. Meghir (2002). "The Effect of School Quality on Educational Attainment and Wages," *Review of Economics and Statistics*, Vol.84(1), 1-20.
- [16] Downes, T. and D. Figlio (1999). "Do Tax and Expenditure Limits Provide a Free Lunch? Evidence on the Link Between Limits and Public Sector Service Quality," *National Tax Journal*, Vol.52(1), 113-28.
- [17] Eherenberg, R.G. and D.J. Brewer (1994). "Do School and Teacher Characteristics Matter? Evidence from High School and Beyond," *Economics of Education Review*, 13, 1-17.
- [18] EU Economic Policy Committee Report (2009). "Efficiency and effectiveness of public expenditure on tertiary education in the EU," *European Economy Occasional Papers* No 70.
- [19] Evans, W.N. and W.E. Oates and R.M. Schwab (1994). "Measuring peer group effects: a study of teenage behavior," *Journal of Political Economy*, 100(5), 966-991.
- [20] Everson, H.T. and R.E. Millsap (2005). "Everyone Gains: Extracurricular Activities in High School and Higher SAT Scores," *College Board Research Report*, 2005-2. New York: College Entrance Examination Board.
- [21] Fan, Y., Q. Li and A. Weersink (1996). "Semi-parametric Estimation of Stochastic Production Frontier Models," *Journal of Business and Economic Statistics*, 14, 460-8.
- [22] Figlio, D.N. (1999). "Functional Form and The Estimated Effects of School Resources," *Economics of Education Review*, 18, 241-252.
- [23] Figlio, D. and S. Loeb, (2011) "School Accountability." In Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, editor: *Handbooks in Economics*, Vol. 3, The Netherlands: North-Holland, pp. 383-421.
- [24] Figlio, D.N. and K.S. Rueben (2001). "Tax limits and the Qualifications of New Teachers" *Journal of Public Economics*, 80, 49C71.

- [25] Figlio, D.N. and A. O’Sullivan (2001). “The Local Response to Tax Limitations Measures: Do Local Governments Manipulate Voters to Increase Revenues?” *Journal of Law and Economics*, 44, 233-257.
- [26] Gong, B.-H. and R.C. Sickles (1992). “A Comparison Between Stochastic Frontier and Data Envelopment Methods Using Panel Data,” *Journal of Econometrics*, 51, 259-84.
- [27] Grosskopf, S., K.J. Hayes, L.T. Taylor and W.L. Weber (2001). “On the Determinants of School District Efficiency: Competition and Monitoring,” *Journal of Urban Economics*, 49, 453-478.
- [28] Gyimah-Brempong, K. and A.O. Gyapong (1991). “Characteristics of Education Production Functions: An Application of Canonical Regression Analysis,” *Economics of Education Review*, 10(1), 7-17.
- [29] Hall, P., Q. Li and J.S. Racine (2007). “Nonparametric Estimation of Regression Functions in the Presence of Irrelevant Regressors,” *Review of Economics and Statistics*, 89, 784-789.
- [30] Hanushek, E.A. (1996). “School resources and students performance, ” In: Burtless, G. (Ed.), *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*. Brookings Institution, Washington, D.C., pp. 43–73.
- [31] Hanushek, E.A. (2003). “The Failure of Input-Based Schooling Policies,” *Economic Journal*, 113, F64-F98.
- [32] Hanushek, E.A. (2010). “Education Production Functions: Developed Countries Evidence,” in Dominic J. Brewer and Patrick J. McEwan (ed.). *Economics of Education* (Amsterdam: Elsevier), 132-136.
- [33] Hanushek, E.A. and M.E. Raymond (2005). “Does School Accountability Lead to Improved Student Performance?” *Journal of Policy Analysis and Management*, Vol. 24, No. 2, 297-327.
- [34] Harris, D.N. (2010). “Education Production Functions: Concepts. ” in *Economics of Education*, edited by D.J. Brewer and P.J. McEwan
- [35] Henderson, D.J. and E. Maasoumi (2014). “Searching for Rehabilitation in Nonparametric Regression Models with Exogenous Treatment Assignment,” in *Oxford Handbook of Applied Nonparametric and Semiparametric Econometrics and Statistics*, Jeffrey S. Racine, Liangjun Su and Aman Ullah (Eds.), 501-520.

- [36] Henderson, D.J., C. Papageorgiou and C.F. Parmeter (2011). "Growth Empirics without Parameters," *Economic Journal*, Vol. 122, 125-154
- [37] Henderson, D.J. and C.F. Parmeter (2015). *Applied Nonparametric Econometrics*, Cambridge University Press, New York
- [38] Henderson, D.J., S. Polachek and L. Wang (2011). "Heterogeneity in Schooling Rates of Return," *Economics of Education Review*, Vol. 30, No. 6, 1202-1214
- [39] Heshmati, A. and S.C. Kumbhakar (1997). "Efficiency of Primary and Secondary Schools in Sweden," *Scandinavian Journal of Educational Research*, Vol. 41, 33-51.
- [40] Holloway, J. (2000). "Extracurricular Activities: The Path to Academic Success?," *Educational Leadership*, 57(4), 87-88.
- [41] Houtenville, A.J. and K.S. Conway (2008). "Parental Effort, School Resources and Student Achievement," *Journal of Human Resources*, Vol. XLIII, No. 2, pp. 437-53.
- [42] Hoxby, C.M. (1999). "The Productivity of Schools and Other Local Public Good Producers," *Journal of Public Economics*, 74, 1-30.
- [43] Imbens, G.W. and J.D. Angrist (1994). "Identification and Estimation of Local Average Treatment Effects," *Econometrica* 62, 467-475.
- [44] Kling, J.R. J.B. Liebman and L.F. Katz (2007). "Experimental Analysis of Neighborhood Effects," *Econometrica* 75, 83-119
- [45] Kneip, A. and L. Simar (1996). "A General Framework for Frontier Estimation with Panel Data," *Journal of Productivity Analysis*, 7, 187-212.
- [46] Kumbhakar, S.C. and C.A.K. Lovell (2000). *Stochastic Frontier Analysis*, Cambridge University Press, New York
- [47] Ladd, H.F. and R.P. Walsh (2002). "Implementing value-added measures of school effectiveness: getting the incentives right," *Economics of Education Review*, 21(1), 1-17.
- [48] Levine, P. B. (2011). How Does Parental Unemployment Affect Children's Educational Performance? In G. J. Duncan & R. Murnane (Eds.), *Whither Opportunity? Rising Inequality and the Uncertain Life Chances of Low-Income Children*. New York: Russell Sage Foundation.
- [49] Li, Q. and J. Racine (2004). "Cross-Validated Local Linear Nonparametric Regression," *Statistica Sinica*, 14, 485-512.

- [50] Li, Q. and J. Racine (2006). *Nonparametric Econometrics: Theory and Practice*, Princeton University Press, Princeton.
- [51] Liu, S. T. (2000). *Essays on the Effects of Immigration on Education and Crime*, Ph.D. thesis, MIT.
- [52] Martins-Filho, C. and F. Yao (2007). “Nonparametric Frontier Estimation via Local Linear Regression, ” *Journal of Econometrics*, 141, 283-319.
- [53] McCarty, T.A. and S. Yaisawarng (1993). “Technical Efficiency in New Jersey School Districts,” in *The Measurement of Productive Efficiency: Techniques and Applications*, edited by H.O. Fried, S.S. Schmidt, and C.A. Knox Lovell. Oxford University Press.
- [54] Millimet, D.L. and T. Collier (2008). “Efficiency in Public Schools: Does Competition Matter?” *Journal of Econometrics*, 145, 134-157.
- [55] Millimet, D.L., J.A. List and T. Stengos (2003). “The Environmental Kuznets Curve: Real Progress or Misspecified Models?” *Review of Economics and Statistics*, 85, 1038-1047.
- [56] Millimet, D.L. and V. Rangaprasad (2007). “Strategic Competition Amongst Public Schools ” *Regional Science and Urban Economics*, 37, 199-219.
- [57] Neymotin, F. “Immigration and its Effect on the College-Going Outcomes of Natives,” *Economics of Education Review* 28, 2009: 538-550.
- [58] Parmeter, C.F. and J. Racine (2012). “Smooth Constrained Frontier Analysis,” in N. Swanson and X. Chen (eds.) *Recent Advances and Future Directions in Causality, Prediction and Specification Analysis* , Springer-Verlag, 463-488
- [59] Park, B., R.C. Sickles and L. Simar (1998). “Stochastic Frontiers: A Semiparametric Approach,” *Journal of Econometrics*, 84, 273-301.
- [60] Park, B., R.C. Sickles and L. Simar (2003). “Semiparametric Efficient Estimation of AR(1) Panel Data Models,” *Journal of Econometrics*, 117, 279-309.
- [61] Park, B. and L. Simar (1994). “Efficient Semiparametric Estimation in a Stochastic Frontier Model,” *Journal of the American Statistical Association*, 89, 929-36.
- [62] Parker, M.N. (2009). “The Present and Future Impact of Tax Caps on Small, Rural Illinois School Districts,” *Unpublished Dissertation*, Western Illinois University. Educational Leadership.

- [63] Racine, J. and Q. Li (2004). "Nonparametric Estimation of Regression Functions with Both Categorical and Continuous Data," *Journal of Econometrics*, 119, 99-130.
- [64] Robinson, P.M. (1988). "Root-n Consistent Semiparametric Regression," *Econometrica*, 56, 931-954.
- [65] Rosthstein, R. (2004). *Class and Schools: Using Social, Economic and Educational Reform to Close the Black-White Achievement Gap*. New York: Teachers College.
- [66] Ruggiero, J. (1996). "Efficiency of Educational Production: An Analysis of New York School Districts," *Review of Economics and Statistics*, 78, 499-509.
- [67] Ruggiero, J., J. Miner and L. Blanchard (2002). "Measuring Equity of Educational Outcomes in the Presence of Inefficiency," *European Journal of Operational Research*, 142, 642-652.
- [68] Schmidt, P., and R.C. Sickles (1984). "Production Frontiers and Panel Data," *Journal of Business and Economic Statistics*, 2, 367-74.
- [69] Sickles, R.C. (2005). "Panel Estimators and the Identification of Firm-Specific Efficiency Levels in Parametric, Semiparametric and Nonparametric Settings," *Journal of Econometrics*, 126, 305-34.
- [70] Smith, P.C. and A. Street (2006). "Analysis of Secondary School Efficiency: Final Report " *U.K. Department for Education and Skills Research Report No 788*.
- [71] Thanassoulis, E. (2000a). "DEA and its use in the regulation of water companies, " *European Journal of Operational Research* 127(1), 1C13.
- [72] Thanassoulis, E. (2000b). "The use of Data Envelopment Analysis in the regulation of UK water utilities: Water distribution, " *European Journal of Operational Research* 126(2), 436C453.
- [73] Todd, P. and K. Wolpin (2003). "On the Specification and Estimation of the Production Function for Cognitive Achievement, " *Economic Journal*, 113(485), F3-F33.
- [74] Vigdor, J. and J. Ludwig (2010). "Neighborhoods and Peers in the Production of Schooling. " in *Economics of Education*, edited by D.J. Brewer and P.J. McEwan.
- [75] Vignoles, A. and R. Levacic and J. Walker and S. Machin and D. Reynolds (2000). "The relationship between resource allocation and pupil attainment: a review. " in *DPUQ, Centre for the Economics of Education, LSE (November)*.

- [76] Wang, H.J. and P. Schmidt (2002). "One-step and Two-step Estimation of the Effects of Exogenous Variables on Technical Efficiency Levels," *Journal of Productivity Analysis*, 18, 129-144

Table 1: LOCAL CONSTANT LEAST SQUARE BANDWIDTHS (VIA LEAST SQUARED CROSS VALIDATION) AND VARIABLE RELEVANCE

Variable	Mixed Data		
	Bandwidth (1)	Bound (2)	Relevant (3)
1 = PTELL in place in county	0.308	0.5	Yes
# schools in district	0.995	1	No
FTE teachers in district	0.993	1	No
LEA administrators in district	1	1	No
current per pupil expenditure	394264.4	0.413	No
capital per pupil exp.	10891594	2.053	No
total population	667149.3	2.338	No
median hh income	271521.4	0.623	No
persons age 20+ w/o HS diploma	633337.4	2.337	No
persons 20+ w/ bachelors degree	0.594	3.059	Yes
Unemployment Rate	0.005	0.052	Yes
% kids, 4-19, speak language other than English	0.033	0.203	Yes
% occupied housing	0.071	0.086	Yes
% owner-occupied housing	0.01	0.194	Yes
Test ID	0.038	0.875	Yes
Year	1	1	No

¹ Notes: For continuous variables, bounds are calculated by two times the standard deviation of the variable. For unordered discrete variables, the upper bound is given by $(c-1)/c$, where c is the number of unique values the unordered categorical variable may take. For ordered discrete variables, the upper bound is unity.

Table 2: LOCAL LINEAR LEAST SQUARE ESTIMATION: BANDWIDTHS (LSCV) AND GRADIENT ESTIMATES

Variable	Bandwidths			Gradient Estimates				
	(1)	(2)	(3)	10 th	25 th	50 th	75 th	90 th
PTELL	0.089	0.500	-0.047*** (0.010)	-0.024* (0.013)	-0.004 (0.003)	0.005 (0.004)	0.022** (0.011)	
% Population with Bachelors Degree	0.821	3.059	-0.019 (0.012)	0.000 (0.018)	0.016* (0.008)	0.029*** (0.007)	0.044** (0.019)	
Unemployment Rate	0.007	0.052	-6.102*** (1.297)	-3.805** (1.707)	-2.062*** (0.660)	-0.728 (0.910)	0.688 (14.077)	
% Non-English Speaking Kids	0.103	0.203	-0.383* (0.195)	-0.136 (0.158)	0.043 (0.132)	0.203** (0.079)	0.405** (0.181)	
% Occupied Housing	0.381	0.086	-1.511*** (0.278)	-0.521* (0.287)	0.067 (0.210)	0.521 (0.790)	1.012** (0.420)	
% Owner-Occupied Housing	0.044	0.194	-0.431* (0.255)	-0.12 (0.146)	0.172 (0.140)	0.463** (0.222)	0.809*** (0.214)	

¹ Notes: For continuous variables, bounds are calculated by 2 times the standard deviation of the variable. For unordered discrete variables, the upper bound is given by $(c-1)/c$, where c is the number of unique values the unordered categorical variable may take. For ordered discrete variables, the upper bound is unity.

² Notes: Bootstrapped standard errors are reported in parentheses.

³ Notes: Asterisks denote statistical significance levels: * * *, ($p \leq 0.01$), **, ($p \leq 0.05$), and *, ($p \leq 0.10$).

Table 3: KOLMOGOROV-SMIRNOV TESTS OF EQUALITY OF THE DISTRIBUTIONS OF PARTIAL EFFECTS BETWEEN DIFFERENT SUBJECT TESTS (P-VALUES)

		Reading			Math			
		Grade 6	Grade 8	Grade 10	Grade 3	Grade 6	Grade 8	Grade 10
Panel A: PTELL								
Reading	Grade 3	0.0336	0.0265	0.0327	0.3056	0.0124	0.0008	0.0025
	Grade 6		0.5781	0.0000	0.0922	0.9906	0.3682	0.0177
	Grade 8			0.0001	0.4653	0.9218	0.5077	0.0372
	Grade 10				0.0006	0.0000	0.0000	0.1567
Math	Grade 3					0.1768	0.0211	0.0032
	Grade 6						0.5630	0.0163
	Grade 8							0.0014
Panel B: % Population with Bachelors Degree								
Reading	Grade 3	0.0000	0.0000	0.0000	0.0005	0.0008	0.0000	0.0000
	Grade 6		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	Grade 8			0.0000	0.0000	0.0000	0.0000	0.0000
	Grade 10				0.0000	0.0000	0.0000	0.0638
Math	Grade 3					0.0388	0.0096	0.0000
	Grade 6						0.0591	0.0000
	Grade 8							0.0000
Panel C: Unemployment Rate								
Reading	Grade 3	0.0000	0.1340	0.0006	0.0000	0.0009	0.0000	0.0000
	Grade 6		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	Grade 8			0.0074	0.0000	0.0003	0.0000	0.0000
	Grade 10				0.0028	0.0030	0.0000	0.0000
Math	Grade 3					0.0003	0.0529	0.0011
	Grade 6						0.0000	0.0303
	Grade 8							0.0007

Table 4: KOLMOGOROV-SMIRNOV TESTS OF EQUALITY OF THE DISTRIBUTIONS OF PARTIAL EFFECTS BETWEEN DIFFERENT SUBJECT TESTS (P-VALUES) - CONT.-

		Reading			Math			
		Grade 6	Grade 8	Grade 10	Grade 3	Grade 6	Grade 8	Grade 10
Panel D: % Non-English Speaking Kids								
Reading	Grade 3	0.0003	0.0841	0.0000	0.0000	0.0046	0.0000	0.1489
	Grade 6		0.0137	0.0000	0.0000	0.0758	0.0000	0.0020
	Grade 8			0.0000	0.0000	0.0015	0.0000	0.2666
	Grade 10				0.0000	0.0000	0.0000	0.0000
Math	Grade 3					0.0006	0.0000	0.0000
	Grade 6						0.0000	0.0140
	Grade 8							0.0000
Panel E: % Occupied Housing								
Reading	Grade 3	0.0000	0.0000	0.0000	0.2461	0.0000	0.0000	0.0000
	Grade 6		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	Grade 8			0.0000	0.0000	0.0000	0.0089	0.0000
	Grade 10				0.0000	0.0000	0.0004	0.0000
Math	Grade 3					0.0000	0.0000	0.0000
	Grade 6						0.0000	0.0043
	Grade 8							0.0000
Panel F: % Owner-Occupied Housing								
Reading	Grade 3	0.0000	0.0000	0.0000	0.0913	0.0000	0.0000	0.4621
	Grade 6		0.0004	0.0001	0.0000	0.0000	0.0099	0.0000
	Grade 8			0.0997	0.0000	0.3712	0.0200	0.0000
	Grade 10				0.0000	0.0359	0.0039	0.0000
Math	Grade 3					0.0000	0.0000	0.0047
	Grade 6						0.0030	0.0000
	Grade 8							0.0007

Table 5: TECHNICAL EFFICIENCIES DESCRIPTIVES

	Minimum (1)	Q1 (2)	Median (3)	Mean (4)	Q3 (5)	Maximum (6)
Nonparametric Production	0.265	0.815	0.835	0.836	0.853	1.000
Cobb-Douglas Production	0.038	0.292	0.413	0.416	0.524	1.000
Translog Production	0.003	0.278	0.414	0.418	0.557	1.000

¹ Notes: $Q1$ and $Q3$ represent the 25th and 75th percentiles of the distribution of technical efficiencies.

Table 6: CORRELATION COEFFICIENTS BETWEEN DIFFERENT MEASURES OF TECHNICAL EFFICIENCIES AND THE LEVEL OF TEST SCORES

Panel A: Spearman				
	Nonparametric	Cobb-Douglas	Translog	Level of Tests
Nonparametric	1.000			
Cobb-Douglas	0.136	1.000		
Translog	0.114	0.981	1.000	
The Level of Tests	0.346	-0.024	-0.034	1.000

Panel B: Kendall				
	Nonparametric	Cobb-Douglas	Translog	Level of Tests
Nonparametric	1.000			
Cobb-Douglas	0.088	1.000		
Translog	0.073	0.893	1.000	
The Level of Tests	0.232	-0.017	-0.024	1.000

Panel C: Pearson				
	Nonparametric	Cobb-Douglas	Translog	Level of Tests
Nonparametric	1.000			
Cobb-Douglas	0.095	1.000		
Translog	0.094	0.971	1.000	
The Level of Tests	0.334	0.008	-0.013	1.000

Table 7: KOLMOGOROV-SMIRNOV TESTS OF THE EFFICIENCY DISTRIBUTIONS

	Equality of Distributions	Stochastic Dominance
	p-values	p-values
	(1)	(2)
Nonparametric v.s. Cobb-Douglas	0.0000	0.9999
Nonparametric v.s. Translog	0.0000	0.9999

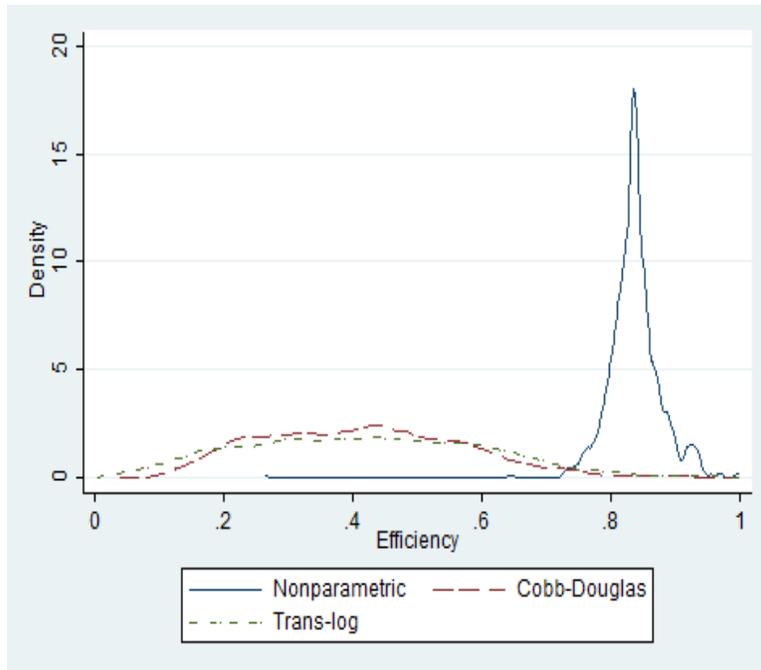
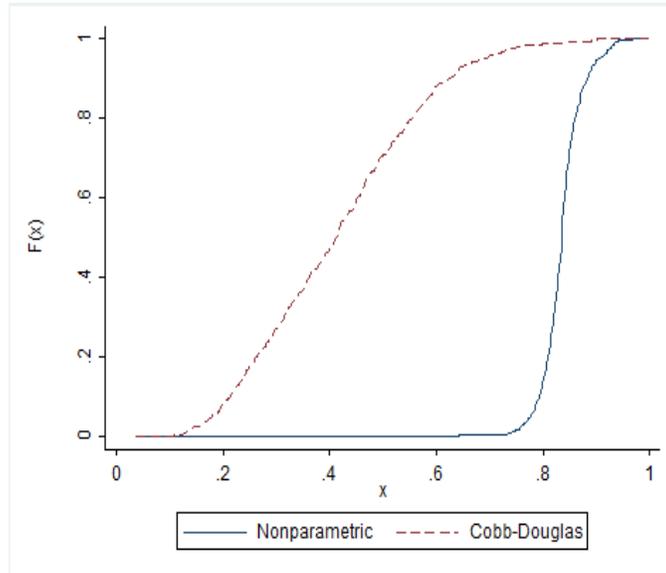
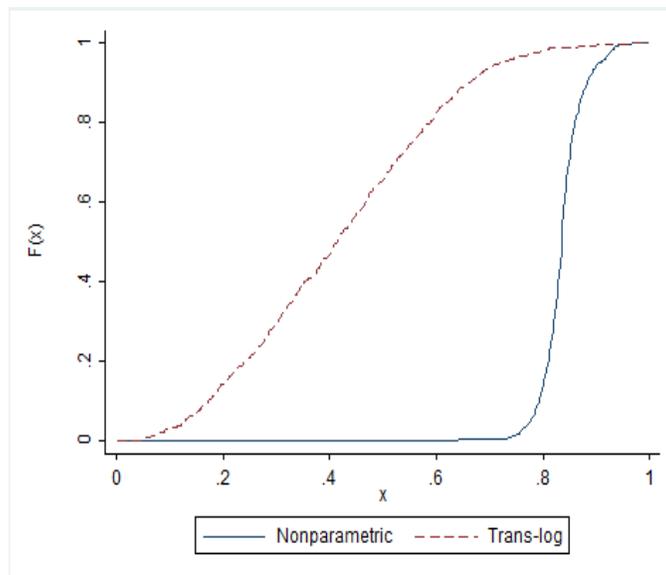


Figure 1: Density of Technical Efficiencies



(a) Nonparametric v.s. Cobb-Douglas



(b) Nonparametric v.s. Translog

Figure 2: Comparisons of Cumulative Distribution Functions

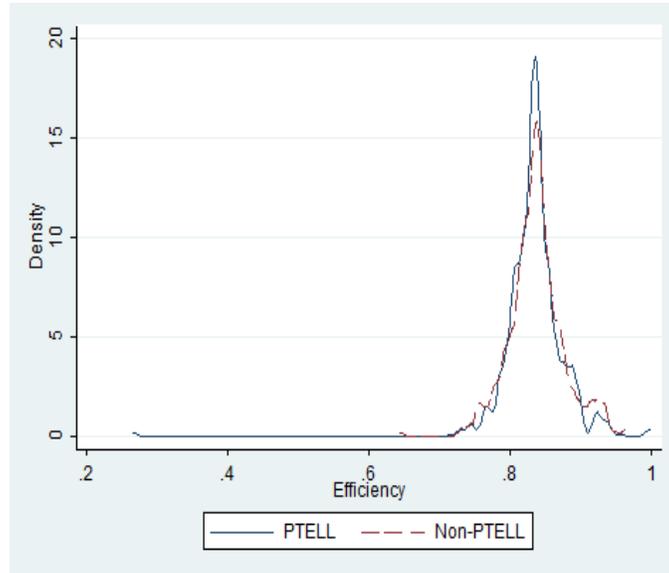


Figure 3: Density of Technical Efficiencies (Nonparametric Results): PTELL v.s. Non-PTELL

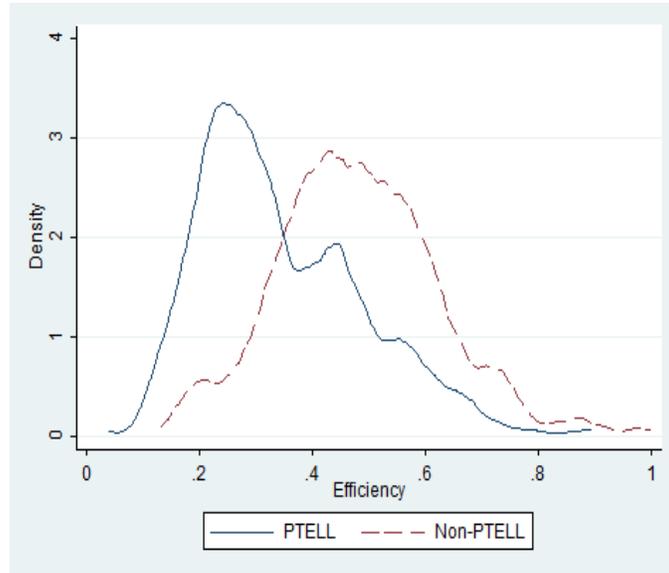


Figure 4: Density of Technical Efficiencies (Cobb-Douglas Results): PTELL v.s. Non-PTELL

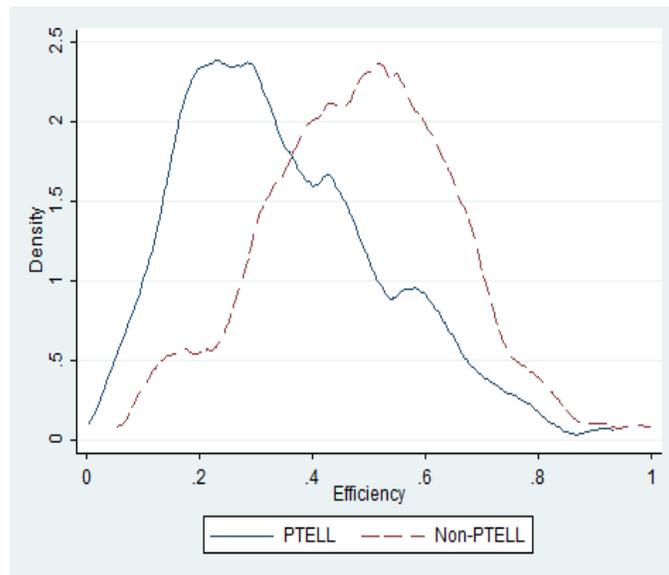


Figure 5: Density of Technical Efficiencies (Translog Results): PTELL v.s. Non-PTELL

7 Appendix

Table 8: SUMMARY STATISTICS

Variable	Mean	S.D
Test Scores	270.168	32.276
1 = PTELL in place in county	0.507	0.500
# schools in district	4.543	20.527
FTE teachers in district	130.638	808.914
LEA administrators in district	3.534	14.210
current per pupil expenditure	5900.962	1670.206
capital per pupil exp.	802.476	1152.165
total population	19036.910	100969.400
median hh income	44880.190	16845.540
persons age 20+ w/o HS diploma	694.183	5558.734
persons 20+ w/ bachelors degree	2955.359	14722.590
Unemployment Rate	0.051	0.026
% kids, 4-19, speak language other than english	0.329	0.104
% occupied housing	0.936	0.041
% owner-occupied housing	0.727	0.099
No. of Observations	1708	

¹ Data Source: Millimet and Collier (2008)

7.1 Kernel Functions for Mixed Continuous and Discrete Data

Recall that we define d_c as the number of continuous inputs. The number of unordered and ordered discrete inputs are defined as d_u and d_o , respectively. Note that the total number of inputs $d = d_c + d_u + d_o$. The generalized kernel is the product of different kernel functions specifically designed for each type as follows (recall that $x_i = [x_i^c, x_i^u, x_i^o] = [(x_{1i}^c, \dots, x_{d_c i}^c), (x_{1i}^u, \dots, x_{d_u i}^u), (x_{1i}^o, \dots, x_{d_o i}^o)]$):

$$K\left(\frac{x_i - x}{h}\right) = \prod_{s=1}^{d_c} k^c(x_{si}^c, x_s^c, h_s^c) \prod_{s=1}^{d_u} k^u(x_{si}^u, x_s^u, h_s^u) \prod_{s=1}^{d_o} k^o(x_{si}^o, x_s^o, h_s^o)$$

where h, λ^u, λ^o are the bandwidths for continuous variables, unordered discrete variables, and ordered discrete variables, respectively. The kernel function for continuous variables is given by

$$k^c(x_{si}^c, x_s^c, h_s^c) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2} \left(\frac{x_i^c - x_s^c}{h_s^c}\right)^2\right\}$$

The kernel function for unordered discrete variables (Aitchison and Aitken, 1976) is given by

$$k^u(x_{si}^u, x_s^u, h_s^u) = \begin{cases} 1 - \lambda_s^u & \text{if } x_{si}^u = x_s^u \\ \frac{\lambda_s^u}{q_s - 1} & \text{otherwise} \end{cases}$$

where q_s is the number of unique values a variable can take. The kernel function for ordered discrete variables is given by:

$$k^o(x_{si}^o, x_s^o, \lambda_s^o) = \binom{q_s}{j} (\lambda_s^o)^j (1 - \lambda_s^o)^{d_s - j} \quad \text{when } |x_{si}^o - x_s^o| = j$$

7.2 Local-Constant Least-Square Estimator

The local-constant least square estimator is given by:

$$\widehat{g}(x) = [i'K(x)i]^{-1} i'K(x)y$$

where i is an $NT \times 1$ vector of ones and $K(x)$ is again a $NJ \times NJ$ diagonal matrix of kernel weighting functions for mixed continuous and discrete data with bandwidth vector $(h, \lambda^u, \lambda^o) = (h_1, \dots, h_d, \lambda_1^u, \dots, \lambda_{d_u}^u, \lambda_1^o, \dots, \lambda_{d_o}^o)$ for the continuous variables and discrete variables, respectively. The kernel functions are the same as described in Section 6.1.