# Estimating Production Functions When Output Prices and Quality Are Unobservable

Xuebing Yang[1]
June 2017

Preliminary Draft

**Abstract**

We develop a method to consistently estimate production functions when output prices and product quality are unobservable. Unobservable output prices and product quality are very likely to be correlated with inputs and can lead to serious bias in the estimates. We show that the markup can serve as a control function for unobserved prices and quality in production functions. The markup can be computed with candidate parameters and data, so our approach does not need more data than the traditional approach. While the traditional approach views revenue as a function of inputs and productivity, our approach views it as a function of inputs, productivity and *markup*, explicitly recognizing the role of prices and quality in determination of revenue. We implement our method as an extension of the proxy-variable approach of estimating production functions pioneered by Olley and Pakes (1996). The empirical results give strong and consistent support to our approach.

*Keywords*: Production Function, Quality, Markup,
*JEL Classification*: D24, L11

[1] Pennsylvania State University at Altoona, Tel: 814-940-3317, email: xyang@psu.edu. I would like to thank the Institute for CyberScience Advanced CyberInfrastructure (ICS-ACI) at Pennsylvania State University for providing the high-performance computing service that made this paper possible.

# 1. Introduction

## 1.1. Description of the Problem

Accurately estimating production functions has remained one of most challenging tasks in empirical economics, principally because many important variables related to the production process are unobservable to the econometrician. Since at least as early as Marschak and Andrews (1944), economists had been grappling with the problem that the productivity is unobservable to economists but is observable to firms and therefore affects their input decisions, a problem that is referred to as the transmission bias (Griliches and Mairesse, 1998). Economists have also discussed the problem of unobservable output prices (e.g., Klette and Griliches, 1996; Mairesse and Jaumandreu, 2005). Recently, economists have started discussing the problem caused by unobservable input prices (e.g., Paul et al., 2016). Unobservable output quality has yet to become a major concern for estimation of production functions, but it may be a problem as well since it is likely to be correlated with other variables in the production function. In this paper, we propose an approach to address unobservable prices and quality of products at the same time. Our approach can be nested in the methods that address the transmission bias.

Unobservable product quality could bias the production function estimates since the omitted quality variable may be correlated with other variables in the production function. In the strict sense, a production function is the mapping between inputs and outputs that are measured in standard units across firms and over time. For output, information on quality is essential to its accurate measurement, but the data are rarely available. When the physical quantity is used as the output measure in the place of quality-adjusted output, both sides of the production function is normalized by the unobservable quality of firm $i$ in period $t$, $Q_{it}$, which is usually subsumed into the error term. Had the variations in quality been uncorrelated with other variables in the production function, the omitted information on quality would be innocuous. Unfortunately, we have ample reasons to believe that the contrary is true. Productive and well-endowed firms are probably more likely to produce higher-quality goods. Product quality is also likely to be intertwined with firms' pricing strategies, further complicating the situation when revenue is used as the output measure. In a brief pilot study, we find that using the physical quantity as output without further corrections would lead to serious symptoms when estimating production functions.

Unobservable prices could bias the production function estimates when revenue is used as the output measure, since the omitted price variable is likely to be correlated with other variables in the production function. When the revenue is used as the output measure in place of quality-adjusted output and the price $P_{it}$ is unobservable, both sides of the production function is multiplied by $P_{it}$ and divided by $Q_{it}$, resulting in two omitted variables, or just one omitted variable if we view them as quality-adjusted price, $P_{it}/Q_{it}$. Economists have sought to purge the effects of unobserved price, but that hardly solve the problem since the model is still subject to the bias arising from unobservable quality. A more systematic approach is needed.

### 1.2. Our Solution

Our solution to the problem of unobserved output prices and quality relies on the observation that a firm has to use more input or better input to produce products of higher quality, keeping its productivity unchanged. It is not a strong assumption per se, but the functional form that we choose to relate unit cost (proxied by the marginal cost $MC_{it}$) and output quality does need more empirical evidence. Fortunately, our empirical results indicate that our functional form assumption is probably not outrageously wrong.

Given our assumption on unit cost and output quality, we show that the markup can serve as a control function for prices and quality, since the unobservable quality-adjusted price, $P_{it}/Q_{it}$, can be replaced by $P_{it}/MC_{it}$ and a term about productivity. The markup of a firm, $P_{it}/MC_{it}$, is usually unobservable, but it can be constructed with production function parameters and data using a methodology advanced by Robert Hall (1986, 1988, 1990). As such, our method does not require more data than the traditional approach does. When an estimator searches through the space of candidate parameters, markups of firms can be computed with candidate parameters and data. The role of prices and quality are explicitly recognized by the model, so the problems associated with unobservable prices and quality are avoided.

To admit firm- and time-specific markups in our models, we follow De Loecker (2011) and De Loecker and Warzynski (2012) to use a translog function. We assume the intermediate goods to be the flexible input and use it to derive markup. The three-factor translog production function makes our estimation job a little challenging, since the dimension of parameter space triples vis-à-vis the common Cobb-Douglas case. However, the increase in the parameter dimensions is necessary for solving the problems at hand. Moreover, the translog production function makes it possible to derive firm-level markups, a variable that is important in many economic inquiries (De Loecker, 2011).

Besides the unobserved prices and quality, we also need to address the classic problem of unobserved productivity when estimating production functions. There has been a vastly rich literature on this problem and an array of solutions have been established. Our solution to the unobservable prices and quality can be nested in any of these solutions. In this work, we choose to implement our method in the setting of the proxy-variable approach first developed by Olley and Pakes (1996).

### 1.3. Two Ways to Understand Our Approach

There are two ways to relate our approach to the traditional approach that uses revenue as the output measure. First, the two approaches have related but distinct theories on how to link revenue to inputs and other variables. The traditional approach views revenue as a function of inputs and technology (i.e., productivity), treating prices and quality as disturbances. **Our approach views revenue as a function of inputs, productivity and markup**, explicitly recognizing the role that prices and quality play in the determination of revenue**.** It holds that inputs and technology (i.e., productivity) generates quality-adjusted output, and that revenue is a function of the quality-adjusted output and a firm's pricing decision, which is proxied by the markup (price/marginal cost). In so doing, the role of prices and product quality in the determination of revenue is rightfully recognized. Ideally, revenue should be treated as a function of quality-adjusted output and the price margin (price/average cost), but it remains unclear to us how to proceed from there. Our empirical results also suggest that the difference between markup and price margin may be negligible.

Second, the two approaches can be viewed as two special cases of a more general econometric model. The more general model is the one where the logarithm of markup is added to a traditional model as an additional explanatory variable. According to our theory, the logarithm of markup serves as the control function for the unobservable prices and quality. The traditional approach does not include this control function, so it can be viewed as a model where the coefficient of the control function is restricted to be zero. Our models predicts that the coefficient on the control function is one, therefore it can be viewed as another restricted version of the nest model. The relation between these models can be used for testing the validity of our models.

### 1.4. Empirical Evidence

The ultimate test of our models would be to compare estimates of traditional approaches, those of our approach, and the "correct" estimates where prices and quality are observed. The problem is that product quality data are usually nonexistent, especially for samples that are large enough to estimate a three-factor translog production function with the proxy-variable approach. As an alternative, we check the returns-to-scale and the markups inferred with the estimated parameters and see if they accord well with textbook theories. We also examine the estimated coefficient on the control function and see if it is close to the value predicted by our theory. Finally, we use the Model and Moment Selection Criteria (MMSC) developed by Andrews and Lu (2001) for GMM estimators to evaluate various models. The empirical results consistently support our proposition that the markup can and should serve as a control function of the unobservable prices and quality in the estimation of production functions.

The multifaceted empirical analysis is conducted in a large array of diverse industries. The 13 industries that we study features various degrees of product homogeneity, market concentration, and technology contents of products. We study the competing approaches both at the industry level and with the pooled sample. All empirical studies give consistent support to our model.

### 1.5. Literature

Our paper is related to the rich literature on transmission bias in estimations of production functions. Griliches and Mairesse (1998) and Ackerberg et al. (2007) provide excellent review on the literature of transmission bias. In light of these reviews, we will only give a brief account of this stream of work.

Over the years, economists have developed methods that rely on fixed effects, instrumental variables, and proxy variables to address the transmission bias. Hoch (1955) proposes the fixed-effect approach, which assumes the unobserved productivity to be predominantly firm-specific and time-invariant. This approach sidesteps the correlation between inputs and productivity by focusing on the within-firm variation in inputs and outputs. Griliches and Mairesse (1998) point out, however, this approach encounters both empirical and theoretical problems. A second approach to address the transmission bias is to use instrumental variables for inputs, or the first differences in inputs. Popular instruments includes factors that may affect input prices, lagged inputs, or other variables. The problem with this approach is that it is difficult to find truly good instruments.

The latest approach to address the transmission bias is developed by Olley and Pakes (1996). In the proxy-variable approach, the unobservable productivity is written as a nonparametric function of inputs and the model is identified by the orthogonality conditions between innovations in productivity and current or lagged input choices. Olley and Pakes use investment as the proxy variable. Levinsohn and Petrin (2003) argue that intermediate goods are a better proxy, since it is less lumpy than investment. Ackerberg et al. (2006) discuss the multicollinearity problem in the first stage estimation of the proxy-variable approach and propose to move the estimation of all parameters to the second stage. Wooldridge (2009) advances a one-step approach, but it is computationally intensive.

Our paper is also related to the literature on unobservable firm-level output prices in the investigations of production activities. The literature has documented significant dispersions in output prices across firms and over time (e.g., Abbott, 1991; Dunne and Roberts, 1992; Roberts and Supina, 1996, 2000; Beaulieu and Mattey, 1999; Bils and Klenow, 2004; Ornaghi, 2006; Foster, Haltiwanger, and Syverson, 2008; Kugler and Verhoogen, 2012). Prior to these studies, however, most researchers have used industry price deflator to construct measures of real output (e.g., Griliches and Mairesse, 1984, Clark and Griliches, 1984, Lichtenberg and Siegel, 1991, Kokkelenberg and Nguyen, 1989). Uniform industry-wise price is an innocuous assumption at the aggregate level, but economists realize that firm-level analysis is sensitive to it. Abbot (1991) first studies the bias that unobservable prices may cause to the estimates of production function and productivity growth equations. Klette and Griliches (1996) address this problem by adding a model of product demand to the estimation.

Our work is also related to the literature on unobservable product quality. Most of these papers discuss how firms signal product quality to consumers (e.g., Rao et al., 1999, Kirmani and Rao, 2000, Shapiro, 1982, Allen, 1984, Milgrom and Roberts, 1986), or how to estimate the product quality (e.g., Hallak and Schott, 2011). These papers shed little light on how to address the unobserved quality when estimating production functions, but their topics attest to the elusiveness of quality information and the its correlation with other variables in the production function, suggesting that a method must be developed to address the omitted-variable problem associated with it in the estimation of production functions.

### 1.6. Outline

The rest of this paper is organized as follows. In the next section, we will use some data to briefly discuss problems of unobserved prices and quality in the estimation of production functions. In Section 3, we will cast the unobservable prices and quality in production functions as an omitted-variable problem and develop a solution to it. In Section 4, we will discuss details regarding the estimation procedure and give the exact forms of some related econometric models. Section 5 gives a brief introduction to the data to be used for our empirical analysis. Section 6 reports the empirical results. Section 7 provides a few remarks on our approach and the results. The last section concludes.

## 2. Motivation

In this section, we will have a brief discussion of the problems related to the missing information on price and quality in estimating production functions.

### 2.1. Correlation between Unobserved Prices and Inputs

The literature has documented significant dispersions in output prices across firms and over time. We find the same pattern in our sample as well. The high degree of variations in the unobserved prices, however, are not necessarily an incorrigible problem in estimating production functions. If the unobserved prices are uncorrelated with inputs, they merely make the estimation less accurate, a problem that can be remedied with a larger sample. Considering the literature on the magnitude of price dispersion and its innocuity in theory, we relegate our findings on the magnitude of output price dispersion to the appendix.

A more serious problem associated with the unobservable prices is that they are very likely to be correlated with inputs. Abbott (1991) finds that output prices are correlated with factor prices in the hydraulic cement industry, which will in all probability affect firms' input decisions. Griliches and Klette (1996) argue that in a static imperfectly competitive environment larger firms tend to charge lower prices. These studies all point to the possibility of inconsistency in the estimation of production functions when prices are unobservable.

In this section we will have a brief look at the correlation between prices and inputs in our data. Our data are on medium or big firms in China during 2000-2007. Details about the data can be found in Section 5.

In Table 1, we report the estimated coefficients when the log price is regressed on the logarithm of each factor for 13 industries and the pooled sample. The underlying econometric model is $\ln(Price_{it}) = \gamma_0 + \gamma_v \ln\left(Factor_{v,it}\right) + \epsilon_{it}$, where $i$, $t$, and $v$ index firms, time periods, and inputs, and $\epsilon_{it}$ is the error term. In total, the table reports results of 42 regressions. The standard errors are reported in the parentheses. The asterisks following the parentheses, $^{***}$, $^{**}$, and $^{*}$, indicate significance levels of 1%, 5%, and 10% respectively. In the last row we report the results based on the pooled sample, where industry fixed effects are included in the econometric model.

The results in Table 1 manifest a very clear and consistent pattern that prices and inputs are strongly correlated. In most cases, prices tend to increase with factor inputs in our samples, suggesting that larger firms tend to charge a higher price. This is not necessarily in contradiction with Griliches and Klette's (1996) assumption, since larger firms may be producing goods of higher quality.

**Table 1: Correlation between Price and Individual Inputs**

$$\ln(Price_{it}) = \gamma_0 + \gamma_v \ln(Factor_{v,it}) + \epsilon_{it}$$

| Product | (1) Capital | (2) Labor | (3) Intermediates | (4) # Obs |
|---|---|---|---|---|
| Bearings | .29 (.03)*** | .31 (.03)*** | .38 (.03)*** | 3471 |
| Cement | .02 (.00)*** | .05 (.00)*** | .09 (.00)*** | 27101 |
| Ceramics | .04 (.02)* | -.29 (.03)*** | .43 (.03)*** | 3514 |
| Engineered Wood | .02 (.01)*** | .10 (.01)*** | .09 (.01)*** | 9493 |
| Ferroalloys | -.00 (.01) | -.01 (.02) | .23 (.01)*** | 4412 |
| Furniture | .12 (.01)*** | -.02 (.02) | .13 (.01)*** | 8134 |
| Garments | .17 (.00)*** | .09 (.01)*** | .32 (.00)*** | 50428 |
| Leather Shoes | .15 (.01)*** | .12 (.01)*** | .28 (.01)*** | 8288 |
| Paperware | .07 (.01)*** | .14 (.01)*** | .17 (.01)*** | 11233 |
| Plastics | .13 (.00)*** | .23 (.01)*** | .19 (.01)*** | 31305 |
| Refractories | .08 (.01)*** | .03 (.02)** | .23 (.01)*** | 4672 |
| Traditional Medicines | .25 (.02)*** | .14 (.02)*** | .34 (.02)*** | 5557 |
| Valves | .25 (.02)*** | .39 (.02)*** | .42 (.02)*** | 3454 |
| All Industries | .12 (.00)*** | .12 (.00)*** | .23 (.00)*** | 171062 |

Table 2 reports the estimation results when the logarithm of prices are regressed on the logarithms of all three factors at the same time. For the last row, industry fixed effects are included in the econometric model. In total, the table reports results of 14 regressions. The format of the statistics in each cell is the same as that in Table 1. There are significant changes in the estimated coefficients compared with Table 1, but the main message remains unchanged: prices are significantly correlated with inputs, and the pattern is observed in all industries studied here.

The above results suggest that a new approach is needed to address the unobservable prices in estimating production functions. When revenue is used as the output measure and prices are unobserved, prices will be subsumed in the error term of the production function. The strong correlation evidenced by Table 1 and Table 2 is very likely to cause bias in the estimates, unless some other unobserved variable (e.g., quality) perfectly offset the variations in prices.

**Table 2: Correlation between Price and All Inputs**

$$\ln(Price_{it}) = \gamma_0 + \gamma_k \ln(Capital_{it}) + \gamma_l \ln(Labor_{it}) + \gamma_m \ln(Intermediates_{it}) + \epsilon_{it}$$

| Product | (1) Capital | (2) Labor | (3) Intermediates | (4) # Obs |
|---|---|---|---|---|
| Bearings | .16 (.04)*** | -.07 (.05) | .28 (.05)*** | 3471 |
| Cement | -.04 (.00)*** | -.01 (.01)* | .13 (.00)*** | 27101 |
| Ceramics | .13 (.02)*** | -.79 (.03)*** | .73 (.03)*** | 3514 |
| Engineered Wood | -.04 (.01)*** | .07 (.02)*** | .10 (.01)*** | 9493 |
| Ferroalloys | -.06 (.01)*** | -.28 (.02)*** | .43 (.02)*** | 4412 |
| Furniture | .16 (.02)*** | -.35 (.02)*** | .23 (.02)*** | 8134 |
| Garments | .13 (.00)*** | -.30 (.01)*** | .38 (.01)*** | 50428 |
| Leather Shoes | .06 (.01)*** | -.31 (.02)*** | .46 (.01)*** | 8288 |
| Paperware | -.04 (.01)*** | .04 (.02)** | .18 (.02)*** | 11233 |
| Plastics | .02 (.01)*** | .14 (.01)*** | .09 (.01)*** | 31305 |
| Refractories | -.00 (.02) | -.19 (.02)*** | .33 (.02)*** | 4672 |
| Traditional Medicines | .17 (.02)*** | -.51 (.04)*** | .49 (.03)*** | 5557 |
| Valves | .02 (.02) | .09 (.03)*** | .35 (.03)*** | 3454 |
| All Industries | .04 (.00)*** | -.12 (.00)*** | .26 (.00)*** | 171062 |

## 2.2. Symptoms Caused by Unobservable Quality

Information on prices is hard to come by, but information on product quality is largely nonexistent. Even if it is available to the econometrician, it remains unclear how to convert the physical output into standard units with the quality information. For instance, it is not immediately clear how to convert a bushel of Six-rowed Blue Malting barley into bushels of Two-rowed barley, suppose the latter is selected as the stand unit.[2] As a result, product quality is almost invariably treated as an unobservable variable in estimations of production functions. In this section, we show that the omitted information on quality may lead to serious symptoms for estimates of production functions.

We estimate production functions using different output measures. We find that the estimates are particularly questionable when the physical quantity is used as the output measure. All estimations that we run are otherwise identical, so the elevated symptoms are likely arising from the measurement errors unique to the physical quantity. Since unobservable quality is a primary source of measurement errors embodied in the physical quantity, we have reasons to believe that the unobservable quality could be a quite damaging factor in the estimation of production functions.

[2] Definitions for barley grades can be found at https://www.gipsa.usda.gov/fgis/standards/810barley97.pdf.

Section 6 will provide a comprehensive analysis of the symptoms. To avoid introducing too many technical details, the brief motivational diagnosis here will be based on the estimated returns to scale and markups. We compute statistics on them for 13 industries and for the pooled sample using the estimates of various production functions. Results using the physical quantity as the output measure are reported in Columns (1), (2), (5) in Table 3. Those using revenue as the output measure are reported in Columns (3), (4), (6) in the same table. The inputs are capital, labor, and intermediate goods. We consider both Cobb-Douglas production functions and translog productions functions. We use the proxy-variable approach to address the transmission bias caused by unobservable productivity, which we will give more details in Section 4.

Columns (1) and (2) show that when the physical quantity is used as the output measure without remedies, the estimated median returns to scale in many industries are far from unit. These figures are hard to believe since they suggest very strong incentives for firms to expand or downsize while we only observe mild changes in firm sizes over time. Columns (5) reveals that many industries have a median markup far below one. This is a serious contradiction with the theory, since the markup is supposed to be at least one. The last row of Table 3 reports the results based on the pooled sample. The results are just as problematic as those at the disaggregated level.

**Table 3: Median Returns to Scale and Markups with Physical Quantity and Revenue as Output**

| Product | Median Returns to Scale | | | | Median Markup | |
|---|---|---|---|---|---|---|
| | (1) Physical Output, Cobb-Douglas | (2) Physical Output, Translog | (3) Revenue, Cobb-Douglas | (4) Revenue, Translog | (5) Physical Output, Translog | (6) Revenue, Translog |
| Bearings | .70 | .71 | 1.01 | .97 | 1.05 | 1.30 |
| Cement | 1.02 | .24 | 1.02 | .99 | 1.49 | 1.27 |
| Ceramics | .82 | 1.02 | 1.03 | .99 | .84 | 1.24 |
| Engineered Wood | 1.16 | .94 | .98 | .99 | 1.15 | 1.27 |
| Ferroalloys | .92 | .96 | .98 | .98 | .63 | 1.28 |
| Furniture | 1.03 | 1.08 | 1.00 | 1.00 | .43 | 1.25 |
| Garments | .87 | .90 | 1.01 | 1.01 | .86 | 1.19 |
| Leather Shoes | .71 | .84 | 1.01 | 1.01 | .44 | 1.22 |
| Paperware | .85 | .92 | 1.02 | 1.00 | 1.01 | 1.23 |
| Plastics | .86 | .86 | 1.00 | 1.00 | 1.17 | 1.20 |
| Refractories | .82 | .87 | 1.02 | 1.01 | .96 | 1.28 |
| Traditional Medicines | 1.20 | .93 | 1.02 | 1.02 | .60 | 1.36 |
| Valves | .82 | .96 | 1.02 | 1.02 | 1.22 | 1.22 |
| All Industries | .67 | .63 | 1.00 | 1.01 | .41 | 1.22 |

Using the physical quantity as the output measure, we also estimate the product functions with other methods to address the unobservable productivity. We try fixed-effects estimator and Arellano-Bond estimator. They both yield returns to scale and markups that are just as doubtable as those observed in Table 3.

Columns (3), (4), and (6) report the results when the revenue is used as the output measure. In these estimations, the unobservables are $P_{it}$ and $1/Q_{it}$, or we can say the unobservable variable is the quality-adjusted price, $P_{it}/Q_{it}$. Columns (3) and (4) show that the estimated median returns to scale in all industries accord well with the theory and existing empirical results. Column (6) shows the estimated median markups also appear to be more reasonable. These results suggest that the omitted-variable problem caused by unobservable $Q_{it}$ may be more seriously than that arises from unobservable $P_{it}/Q_{it}$.

The above analysis provides evidence that prices are correlated with inputs, and that unobservable quality can cause serious problems for the estimation of production functions. The symptoms caused by unobservable quality discussed above alleviate when we let the unobservable quality and prices offset each other, but we are agnostic about to what extent the omitted-variable problem is cured. In the next section, we develop a framework that will address the problems caused by the both missing variables.

## 3. Omitted-Variable Problem and Our Solution

In this section we provide a theoretical analysis of the problems arising from unobservable prices and quality in the omitted-variable framework. We propose a solution based on an assumption about the relation between quality and per unit cost. Given this assumption, the revenue can be viewed as a function of inputs, productivity, and markup. The new econometric model recognizes the role of prices and quality in the determination of revenue, but does not require more data than the traditional approach does.

### 3.1. Description of the Problem

A production function relates the output of a production process to the inputs. When it is used to analyze production activities, the output of different firms should be measured on the same footing. In this work, we assume that quality-adjusted output is a proper standard measure of output. Suppose the quality-adjusted output of firm $i$ in period $t$, $\tilde{Y}_{it}$, is given by the following function

$$\tilde{Y}_{it} = F(X_{it}, \beta) \cdot \exp(\omega_{it} + \epsilon_{it}), \tag{1}$$

where $X_{it}$ is a vector of inputs, $\omega_{it}$ is the productivity observed by the firm and thus affects the firm's choice of inputs, $\epsilon_{it}$ is the shock to firm's productivity that is not observed by the firm, and $\beta$ is a vector of parameters common to all firms.

Equation (1) is the theoretic starting point of analyzing production activities, but it has little empirical relevance since quality-adjusted output, $\tilde{Y}_{it}$, is rarely observable. The physical quantity of output, which we denote with $Y_{it}$, however, may be observable in some fortunate occasions. Suppose the quality-adjusted output equals $Y_{it}$ times a quality measure, $Q_{it}$, then production function (1) can be written as

$$Y_{it} = F(X_{it}, \beta) \cdot \exp(\omega_{it} + \epsilon_{it})/Q_{it}. \tag{2}$$

There are various approaches to estimate Equation (2), depending on our assumptions regarding $\omega_{it}$, $\epsilon_{it}$, and the unobservable $Q_{it}$.

As it was mentioned above, the physical quantities of output are seldom available to the econometrician. Revenue, which we denote with $R_{it}$, is usually the only observable output measure. Suppose the price is $P_{it}$, then the output, measured as revenue, is equal to

$$R_{it} = F(X_{it}, \beta) \cdot \exp(\omega_{it} + \epsilon_{it}) \cdot P_{it}/Q_{it}. \tag{3}$$

Equation (3) is the model that has been frequently estimated to study topics related to production. The difficulty of consistently estimating the model lies in the fact that its right-hand side have three unobservable variables that are very likely to vary with the inputs: $\omega_{it}$, $P_{it}$, $Q_{it}$. There has been a rich literature on how to handle the problem caused by the unobservable $\omega_{it}$, which is termed as the transmission bias, but there has been much less discussion on the unobservable $P_{it}$ and $Q_{it}$. In this paper, we attempt to bridge this gap.

Since the problem of unobservable prices and quality is similar to that of unobservable productivity, we might want to draw inspirations from the solutions that economists have developed for the transmission bias. As it was mentioned in the introduction, over the years, economists have tried fixed-effects, instrumental variables, and proxy-variable approaches to this end. The fixed-effect approach probably would not work very well in our case, since prices and quality are very likely to evolve over time. The instrumental-variable approach is appealing in theory, but it would be difficult to find good instruments for firm-level prices and quality. Any variable that is correlated with prices or quality is most likely to be correlated with inputs. Proxy-variable approaches would require us to develop theories on how prices and quality evolve over time and moment conditions, which would be a massive undertaking but probably a promising direction for future research. In this paper, we propose an approach that is similar to the proxy-variable approach but less complicated. Like the proxy-variable approach, our approach also involves a control function of the unobservables. Unlike the proxy-variable approach, our method does not rely on assumptions on the evolutions of the unobservables to achieve identification. Instead, the functional form assumption dictate the coefficient on the extra term to be one and subsume an unidentified parameter in productivity. In the next section we present our assumption.

### 3.2. Assumption on Quality and Marginal Cost

It usually takes firms better or more inputs to produce products of higher quality, keeping their technology and the amount of output unchanged. For example, if the tire company Michelin wants to produce UTQG Grade AA tires instead of Grade A tires, it will most likely have to use better rubber and fabric, which will in all probability cost the company more money to produce a tire. As another example, if a publisher wants to produce manuscripts with fewer typos, it will probably have to ask its proofreaders to spend more time checking the manuscripts, which will lead to higher cost per manuscript.

The positive correlation between quality and unit cost is self-evident, but it is probably difficult to defend any specific function that relates them. As a matter of fact, we simply choose a function that can most conveniently operationalize the above intuition. We use the marginal cost to represent the unit cost, and assume quality to be a product of marginal cost and productivity:

*Assumption: A firm's product quality in period* $t$ *is given by* $Q_{it} = MC_{it} \cdot \exp\left(\zeta(\omega_{it} + \epsilon_{it})\right)$, *where* $\zeta$ *is a positive constant.*

Note that the marginal cost, $MC_{it}$, is defined as the increase in cost incurred by an additional *physical* unit of product. Given a firm's productivity, $\omega_{it} + \epsilon_{it}$, the above assumption states that the product quality would be higher if the firm spends more money on a unit of output. If the firm becomes more productive, $MC_{it}$ would decrease proportionally, *ceteris paribus.* However, if the firm decides to spend the same amount of money on each unit of product as before, the product quality will increase.

The main issue with the above assumption is that in a given production run the marginal cost varies with the output level, while the product quality is believed to be predetermined and invariant to the output level. The severity of the problem depends on how much the marginal cost varies from unit to unit. Our empirical results find that almost all firms in our sample have strict constant returns to scale. For example, estimated returns to scale of all firms of all time periods have a mean of 1.006 and a standard deviation of 0.005.[3] Adjustment costs will cause marginal costs to deviate from average cost, but the estimated returns to scale give us some reassurance for using marginal cost as a proxy of unit cost.

### 3.3. Revenue as a Function of Markup, Inputs and Productivity

Given our assumption concerning quality, Equation (3) can be written in the following form:

$$r_{it} = \ln \mu_{it} + f(x_{it}, \beta) + (1 - \zeta)(\omega_{it} + \epsilon_{it}), \tag{4}$$

where $r_{it}$=$\ln R_{it}$, $\mu_{it} = P_{it}/MC_{it}$ is the markup, $x_{it} = \ln X_{it}$, and $f(x_{it}, \beta) = \ln F(X_{it}, \beta)$. Equation (4) illustrates the main proposition of our method: *revenue should be view as a function of inputs, productivity, and markup.* The markup in the model explicitly recognizes the role of prices and quality in the determination of revenue, avoiding the problem they may cause when they are subsumed into the error term.

In Appendix B, based on the theory developed by Hall (1986, 1988, 1990) and extended by De Loecker (2011), and De Loecker and Warzynski (2012), we show that under very weak assumptions, the markup of a firm equals

$$\mu_{it} = \theta_{it}^{v}/a_{it}^{v}, \tag{5}$$

where $\theta_{it}^{v}$ is the physical output elasticity of input $v$, and $a_{it}^{v}$ is the expenditure on input $v$ as a share of the firm's revenue. Input $v$ in Equation (5) should be a flexible input. In this work, we assume the intermediate good is flexible.

Given production function (2), the physical output elasticity of intermediate good is

$$\theta_{it}^{m}(x_{it}, \beta) = f_m(x_{it}, \beta) - \frac{\partial \ln Q_{it}}{\partial m_{it}}. \tag{6}$$

Equation (6) shows that the output elasticity of intermediate good consists of its effect on the quality adjusted output and its effect on the product quality. Incidentally, when changes in $M_{it}$

[3] The results are a more accurate version of the one in the last row of Column (2) in Table 10.

causes in an increase in quality, a given amount of quality-adjusted products will be counted as fewer units of physical product, hence $\frac{\partial \ln Q_{it}}{\partial m_{it}}$ enters Equation (6) with a negative sign. Given the output elasticity of the intermediate good, the revenue can be written as

$$r_{it} = \ln\left(\theta_{it}^{m}(x_{it}, \beta)/a_{it}^{m}\right) + f(x_{it}, \beta) + (1-\zeta)(\omega_{it} + \epsilon_{it}). \tag{7}$$

Let $\widetilde{\omega}_{it} = (1-\zeta)\omega_{it}$ and $\tilde{\epsilon}_{it} = (1-\zeta)\epsilon_{it}$, then we have:

$$r_{it} = \ln\left(\theta_{it}^{m}(x_{it}, \beta)/a_{it}^{m}\right) + f(x_{it}, \beta) + \widetilde{\omega}_{it} + \tilde{\epsilon}_{it}. \tag{8}$$

Once we parameterize $f(x_{it}, \beta)$ we can write $\theta_{it}^{m}(x_{it}, \beta)$ as a function of $x_{it}$ and $\beta$. We assume the inputs include of capital, $K_{it}$, labor, $L_{it}$, and intermediate good, $M_{it}$. We assume that $f(x_{it}, \beta)$ takes the translog form:

$$\begin{aligned} f(x_{it}, \beta) = {} & \beta_k k_{it} + \beta_l l_{it} + \beta_m m_{it} + \beta_{kk} k_{it}^2 + \beta_{ll} l_{it}^2 + \beta_{mm} m_{it}^2 + \beta_{kl} k_{it} l_{it} + \beta_{km} k_{it} m_{it} \\ & + \beta_{lm} l_{it} m_{it}. \end{aligned} \tag{9}$$

where $k_{it}$, $l_{it}$, and $m_{it}$ are the logarithms of $K_{it}$, $L_{it}$, and $M_{it}$. Given the translog production function, the term $f_m(x_{it}, \beta)$ in Equation (6) is

$$f_m(x_{it}, \beta) = \beta_m + 2\beta_{mm} m_{it} + \beta_{km} k_{it} + \beta_{lm} l_{it}. \tag{10}$$

In Appendix C, we show that

$$\frac{\partial \ln Q_{it}}{\partial m_{it}} = -\frac{2\beta_{mm}}{f_m(x_{it}, \beta)} - f_m(x_{it}, \beta) + 1. \tag{11}$$

Plugging Equations (9), (10), and (11) in Equation (8), we will be able to write $r_{it}$ as a long function of $x_{it}$, $\beta$, and $\zeta_{it}$

$$\begin{aligned} r_{it} = \ln\Big( & \big(2(\beta_m + 2\beta_{mm} m_{it} + \beta_{km} k_{it} + \beta_{lm} l_{it}) \\ & + 2\beta_{mm}/(\beta_m + 2\beta_{mm} m_{it} + \beta_{km} k_{it} + \beta_{lm} l_{it}) - 1\big)/a_{it}^{m}\Big) + \beta_k k_{it} + \beta_l l_{it} \\ & + \beta_m m_{it} + \beta_{kk} k_{it}^2 + \beta_{ll} l_{it}^2 + \beta_{mm} m_{it}^2 + \beta_{kl} k_{it} l_{it} + \beta_{km} k_{it} m_{it} + \beta_{lm} l_{it} m_{it} \\ & + \widetilde{\omega}_{it} + \tilde{\epsilon}_{it}. \end{aligned} \tag{12}$$

On the right-side hand of the above equation, the first term is the logarithm of the markup, $\ln \mu_{it}$. The next nine terms are $f(x_{it}, \beta)$. The last two terms are productivity terms.

A remark on $\frac{\partial \ln Q_{it}}{\partial m_{it}}$ is in order. It is the term we calculate in Equation (11) and insert in the final equation. The term measures how the quality varies with the flexible input. In developing our assumption about quality and unit cost, however, we have in our mind a world where a firm's product quality remains unchanged in a given period. To investigate the discrepancies, we consider an equation where $\frac{\partial \ln Q_{it}}{\partial m_{it}}$ is restricted to be zero. With this constraint, Equation (8) becomes

$$
\begin{aligned}
r_{it} = \ln\left((\beta_m + 2\beta_{mm} m_{it} + \beta_{km} k_{it} + \beta_{lm} l_{it})/a_{it}^m\right) + \beta_k k_{it} + \beta_l l_{it} + \beta_m m_{it} + \beta_{kk} k_{it}^2 \\
+ \beta_{ll} l_{it}^2 + \beta_{mm} m_{it}^2 + \beta_{kl} k_{it} l_{it} + \beta_{km} k_{it} m_{it} + \beta_{lm} l_{it} m_{it} + \widetilde{\omega}_{it} + \tilde{\epsilon}_{it}.
\end{aligned}
\tag{13}
$$

On the right-side hand of the above equation, the first term is the logarithm of the markup, $\ln \mu_{it}$. We refer to the new markup as the corrected markup.

### 3.4. Bias Due to Unobservable Prices and Qualities

It is of interest to have a theoretical analysis of the bias in estimated coefficients due to unobservable prices and quality. In order to focus on this question, let us temporarily assume $\omega_{it}$ to be observable. We also assume $\frac{\partial \ln Q_{it}}{\partial m_{it}} = 0$, since it will greatly simplify our analysis. With these assumptions, the population model is Equation (13), with the change that the term $\widetilde{\omega}_{it}$ is now observable up to a scale (since $\zeta$ is unobserved). We use a second order Taylor series to approximate the logarithm of the markup, $\ln((\beta_m + 2\beta_{mm} m_{it} + \beta_{km} k_{it} + \beta_{lm} l_{it})/a_{it}^m)$, and relegate the residual to the error term $\tilde{\epsilon}_{it}$. We expand $\ln \mu$ around 1 since it is convenient and in most cases our estimated markups are not very far from 1. Thus the population model can be written as:

$$
\begin{aligned}
r_{it} = (\beta_m + 2\beta_{mm} m_{it} + \beta_{km} k_{it} + \beta_{lm} l_{it})/a_{it}^m - 1 \\
- \frac{1}{2}\left((\beta_m + 2\beta_{mm} m_{it} + \beta_{km} k_{it} + \beta_{lm} l_{it})/a_{it}^m - 1\right)^2 + \beta_k k_{it} + \beta_l l_{it} + \beta_m m_{it} \\
+ \beta_{kk} k_{it}^2 + \beta_{ll} l_{it}^2 + \beta_{mm} m_{it}^2 + \beta_{kl} k_{it} l_{it} + \beta_{km} k_{it} m_{it} + \beta_{lm} l_{it} m_{it} + \widetilde{\omega}_{it} + \tilde{\epsilon}_{it}.
\end{aligned}
\tag{14}
$$

The expanded form of Equation (14) is quite long, so we will not report it here. Let us use $\dot{\beta}_x$ to denote the coefficient on input $x$ in Equation (14). We have

$$\dot{\beta}_k = \beta_k + 2\beta_{km}/a_{it}^m - \beta_m\beta_{km}/(a_{it}^m)^2, \tag{15}$$
$$\dot{\beta}_l = \beta_l + 2\beta_{lm}/a_{it}^m - \beta_m\beta_{lm}/(a_{it}^m)^2, \tag{16}$$
$$\dot{\beta}_m = \beta_m + 4\beta_{mm}/a_{it}^m - 2\beta_m\beta_{mm}/(a_{it}^m)^2, \tag{17}$$
$$\dot{\beta}_{kk} = \beta_{kk} - \beta_{km}^2/(a_{it}^m)^2, \tag{18}$$
$$\dot{\beta}_{ll} = \beta_{ll} - \beta_{lm}^2/(a_{it}^m)^2, \tag{19}$$
$$\dot{\beta}_{mm} = \beta_{mm} - 2\beta_{mm}^2/(a_{it}^m)^2, \tag{20}$$
$$\dot{\beta}_{kl} = \beta_{kl} - \beta_{km}\beta_{lm}/(a_{it}^m)^2, \tag{21}$$
$$\dot{\beta}_{km} = \beta_{km} - 2\beta_{mm}\beta_{km}/(a_{it}^m)^2, \tag{22}$$
$$\dot{\beta}_{lm} = \beta_{lm} - 2\beta_{mm}\beta_{lm}/(a_{it}^m)^2. \tag{23}$$

Since $a_{it}^m$ varies across firms and time, simple analytical forms of the bias in the coefficients generally do not exist. Nonetheless, it is clear from Equations (15)-(23) that all coefficients will be biased.

To put the degree of the bias in perspective, let us consider a back-of-envelope calculation. Our empirical exercise shows that the estimated coefficients of first-order terms in the translog function ($\beta_k$, $\beta_l$, and $\beta_m$) are on the order of 0.1 and those of second-order terms (e.g., $\beta_{kk}, \beta_{kl}, \ldots$) are on the order of 0.01. In our sample, $a_{it}^m$ is about 0.75 on average. In a case where $a_{it}^m$ is 0.75 for all firms, the bias in each estimated coefficients is probably a few percent to ten percent of the parameters' true values. Take $\beta_{lm}$ for instance. The percentage bias in it is $\left(\beta_{lm} - \dot{\beta}_{lm}\right)/\beta_{lm} = 2\beta_{mm}/(a_{it}^m)^2 = 3.56\beta_{mm}$, which is probably a few percent.

## 4. Estimation of Our Models and Related Models

This section explains how to estimate our models with the proxy-variable approach. In terms of coding, our method only makes a few simple changes to the proxy-variable approach. However, since the proxy-variable is a new and evolving method, we plan to give the details of our estimation procedure.

### 4.1. The Proxy-Variable Approach

We will use our workhorse models, Equations (12) and (13), to explain how the estimation procedure works. The procedure is largely the same for the estimation of some related models that we will introduce in the next section.

The proxy-variable approach is one of the econometric methods developed to address the classical transmission bias caused by $\tilde{\omega}_{it}$. Parameter $\tilde{\omega}_{it}$ is unobservable to the econometrician but is observable to firms and will most likely affect the amounts of inputs, $X_{it}$. The main idea of the proxy-variable approach is to assume (1) that there exists an input that monotonically increases with current-period productivity, such as contemporary investment or intermediate goods, (2) that the innovations in the productivity are orthogonal to input decisions that are made before the productivity is revealed to the firm.

Ackerberg et al. (2006) refine the two-stage method developed by Olley and Pakes (1996) and Levinsohn and Petrin (2003). Wooldridge's (2009) one-step approach more efficient and can generate standard errors of estimates, but it is much more computationally intensive. In this work, we adopt Ackerberg, Caves, and Frazer's approach. In what follows, we discuss the details of the procedure.

**First Stage: Estimating Unobserved Productivity Shock $\epsilon_{it}$**

The proxy approach is based on the monotonic functional relationship between inputs and the unobservable productivity. Following this line of literature, we assume that $\widetilde{\omega}_{it}$ can be written as a function of inputs, $\widetilde{\omega}_{it} = \widetilde{\omega}(l_{it}, k_{it}, m_{it})$. In our exercise, we assume that $\widetilde{\omega}(\cdot)$ is a fourth-order polynomial. We assume $\ln \mu_{it}$ can also be written as a fourth-order polynomial function of inputs. Then equations based on Equation (8) can be written as

$$r_{it} = \phi(k_{it}, l_{it}, m_{it}) + \tilde{\epsilon}_{it}. \tag{24}$$

where $\phi(k_{it}, l_{it}, m_{it})$ is defined as:

$$\phi(k_{it}, l_{it}, m_{it}) = \ln\left(\theta_{it}^{m}(x_{it}, \beta)/a_{it}^{m}\right) + \ln f(X_{it}, \beta) + \widetilde{\omega} = \sum_{f=0}^{4} \sum_{g=0}^{4-f} \sum_{h=0}^{4-f-g} \gamma_{fgh} k_{it}^{f} l_{it}^{g} m_{it}^{h}. \tag{25}$$

In the above equation, $\gamma_{fgh}$ is a coefficient on $k_{it}^{f} l_{it}^{g} m_{it}^{h}$, where $f$, $g$, and $h$ are exponents.

In the first stage of our estimation, we estimate $r_{it} = \phi(k_{it}, l_{it}, m_{it}) + \tilde{\epsilon}_{it}$ and store the predicted values $\hat{\phi}_{it}$ for later use.

**Second Stage: Searching Parameter Space**

In the second stage, we iterate through candidate production function parameters $\beta'$ and look for the ones that minimize the criterion function defined by the moment conditions, which will be introduced momentarily.

Given the candidate parameters $\beta' = (\beta_l', \beta_k', \beta_m' , \beta_{kk}', \beta_{ll}', \beta_{mm}', \beta_{kl}', \beta_{km}', \beta_{lm}')$, we can construct the implied value for $\widetilde{\omega}_{it}$ (refer to Equation (24)):

$$\widehat{\widetilde{\omega}}_{it}(\beta) = \widehat{\phi}_{it} - \ln(\theta_{it}^m(x_{it}, \beta')/a_{it}^m) - f(x_{it}, \beta'). \tag{26}$$

If we regress $\widehat{\widetilde{\omega}}_{it}$ on $\widehat{\widetilde{\omega}}_{it-1}$ and other variables that may affect productivity, the residuals, $\hat{\xi}_{it}\ (\beta)$, will be the innovations in the productivity shocks.

The production function parameters are identified with the moment conditions that the innovation in productivity in period $t$, $\xi_{it}(\beta)$, is orthogonal to the fixed input in period $t$, $k_{it}$, the lagged value of the variable inputs, and the interactions between these variables. The orthogonality conditions are based on the timing of input decisions and productivity evolutions. More details on the theory behind these moment conditions can be found in Olley and Pakes (1996), Levinsohn and Petrin (2003), and Ackerberg et al. (2006). Our three-factor translog model has at least nine parameters, so we will need a relatively large set of moment conditions. Let $z_{it}$ be a vector of variables that are supposed to be orthogonal to the innovations in productivity:

$$\begin{gathered} z_{it} = (1, k_{it}, k_{it-1}, l_{it-1}, m_{it-1}, k_{it}^2, , k_{it-1}^2, l_{it-1}^2, m_{it-1}^2, k_{it}k_{it-1}, \\ k_{it}l_{it-1}, k_{it}m_{it-1}, k_{it-1}l_{it-1}, k_{it-1}m_{it-1}, l_{it-1}m_{it-1})' . \end{gathered} \tag{27}$$

The moment conditions that will help us identify the parameters are

$$E\big(z_{it}\xi_{it}(\beta)\big) = 0. \tag{28}$$

Since we have more moment conditions than the parameters, we have over-identifying conditions that can be used to test our model.

The criterion function of the GMM estimator is

$$\sum_{h=1}^{15} \left[ \left( \sum_i \sum_t \xi_{it}(\beta) z_{i,ht} \right)^2 \right], \tag{29}$$

where $z_{i,ht}$ is the $h^{th}$ element of $z_{it}$. Now we can use a program to search the optimal $\beta$ that minimizes criterion function (29). The resulting $\hat{\beta}$ would be our estimate for the production function parameters.

### 4.2. Related Econometric Models

Equations (12) and (13) represents the new method that we propose. Besides them, we will consider some other models. We will give the exact specifications of them in this section for clarity.

Among these alternative models, four of them are traditional Cobb-Douglas or translog models that use either the physical quantity or revenue as the output measure:

- Cobb-Douglas function with physical output as the dependent variable:

$$y_{it} = \beta_k k_{it} + \beta_l l_{it} + \beta_m m_{it} + \omega_{it} + \epsilon_{it}. \tag{30}$$

- Cobb-Douglas function with revenue as the dependent variable:

$$r_{it} = \beta_k k_{it} + \beta_l l_{it} + \beta_m m_{it} + \omega_{it} + \epsilon_{it}. \tag{31}$$

- Translog function with physical output as the dependent variable:

$$\begin{aligned} y_{it} = {} & \beta_k k_{it} + \beta_l l_{it} + \beta_m m_{it} + \beta_{kk} k_{it}^2 + \beta_{ll} l_{it}^2 + \beta_{mm} m_{it}^2 + \beta_{kl} k_{it} l_{it} + \beta_{km} k_{it} m_{it} \\ & + \beta_{lm} l_{it} m_{it} + \omega_{it} + \epsilon_{it}. \end{aligned} \tag{32}$$

- Translog function with revenue as the dependent variable:

$$\begin{aligned} r_{it} = {} & \beta_k k_{it} + \beta_l l_{it} + \beta_m m_{it} + \beta_{kk} k_{it}^2 + \beta_{ll} l_{it}^2 + \beta_{mm} m_{it}^2 + \beta_{kl} k_{it} l_{it} + \beta_{km} k_{it} m_{it} \\ & + \beta_{lm} l_{it} m_{it} + \epsilon_{it}. \end{aligned} \tag{33}$$

Besides these traditional models, we also consider extensions of our workhorse models. Our theory predicts that the coefficient of the control function, $\ln \mu$, is one. We also consider models in which the coefficient on $\ln \mu$, denoted with $\beta_{Markup}$, is a free parameter:

$$\begin{aligned} r_{it} = {} & \beta_{Markup} \ln \Big( \big( 2(\beta_m + 2\beta_{mm} m_{it} + \beta_{km} k_{it} + \beta_{lm} l_{it}) \\ & + 2\beta_{mm} / (\beta_m + 2\beta_{mm} m_{it} + \beta_{km} k_{it} + \beta_{lm} l_{it}) - 1 \big) / a_{it}^m \Big) + \beta_k k_{it} + \beta_l l_{it} \\ & + \beta_m m_{it} + \beta_{kk} k_{it}^2 + \beta_{ll} l_{it}^2 + \beta_{mm} m_{it}^2 + \beta_{kl} k_{it} l_{it} + \beta_{km} k_{it} m_{it} + \beta_{lm} l_{it} m_{it} \\ & + \widetilde{\omega}_{it} + \tilde{\epsilon}_{it}. \end{aligned} \tag{34}$$

$$\begin{aligned} r_{it} = {} & \beta_{Markup} \ln \big( (\beta_m + 2\beta_{mm} m_{it} + \beta_{km} k_{it} + \beta_{lm} l_{it}) / a_{it}^m \big) + \beta_k k_{it} + \beta_l l_{it} + \beta_m m_{it} \\ & + \beta_{kk} k_{it}^2 + \beta_{ll} l_{it}^2 + \beta_{mm} m_{it}^2 + \beta_{kl} k_{it} l_{it} + \beta_{km} k_{it} m_{it} + \beta_{lm} l_{it} m_{it} + \widetilde{\omega}_{it} + \tilde{\epsilon}_{it}. \end{aligned} \tag{35}$$

Equations (34) and (35) encompass the traditional approach (represented by Equation (33) ) and our approach as two special cases.

## 5. Data

Out empirical analysis is based on the China Annual Survey of Industrial Firms and the China Industrial Output Database. Both are annual survey databases published by the National Bureau of Statistics of China. The former is more well-known in academia outside of China.[4] The databases include all firms with sales over 5 million Chinese Yuan (about 600,000 US dollars). Firms in the databases account for about 90 percent of output of the covered industries. The Annual Survey of Industrial Firms reports various administrative and financial facts of firms. The Output Database only has information on quantities of major outputs. The two databases have an overlapping coverage period from 2000 to 2007.

The unit of observation in the databases is enterprise. About 98% of manufacturing firms have only one unit, which is a subdivision that usually focuses on a group of closely related products. On average each manufacturing firm has 1.05 units. Each firm has an organization registration code issued by the government. We use this code to identify firms.

We discount all the monetary values in the Annual Survey of Industrial Firms with China's GDP deflator. We delete observations that report very implausible growth in major variables. We also delete the 1-percent outliers in terms of input-input or input-output ratios.

The Output Database records the annual information on the major products of firms. Due to economies of scope, multiple products are often produced at the same firm. For example, among firms that produce Wall-mounted Gas Water Heaters, 82.9 percent of them also produce Gas Ranges. Some of the products produced by the same firm may be very different in nature. For example, 58.6 percent of Sugar producers also produce electricity, presumably using bagasse – the fibrous residue left after sugarcane is crushed – as fuel. Production of multiple products at the same firm makes it difficult to calculate product prices with data at hand, since we only have the yearly total value of all products for each firm. Therefore, we try to focus on products that seldom have "siblings". In these industries, we focus on observations where the firm only reports one product in the year.

[4] For more details regarding the The Annual Survey of Industrial Firms, please refer to Song et al. (2011) and Hsieh and Klenow (2009).

**Table 4: Sibling Shares of Selected Industries**

| Product | (1) # Obs in Output Data | (2) # Producer | (3) % Multi-product | (4) Sibling 1 | (5) Sibling 2 |
|---|---|---|---|---|---|
| Bearings | 4268 | 1111 | .04 | Bricks (.01) | Plastics (.01) |
| Cement | 36697 | 9193 | .09 | Electricity (.04) | Coal (.03) |
| Ceramics | 4682 | 1329 | .06 | Refractories (.02) | Cement (.01) |
| Engineered Wood | 12435 | 4164 | .08 | Veneer (.02) | Furniture (.02) |
| Ferroalloys | 6404 | 2275 | .15 | Electricity (.06) | Calcium Carbide (.03) |
| Furniture | 10748 | 3439 | .10 | Engineered Wood (.02) | Plastics (.02) |
| Garments | 65494 | 18543 | .09 | Knitted Products (.03) | Fabrics (.02) |
| Leather Shoes | 10644 | 2950 | .06 | Garments (.02) | Rubber Boots (.01) |
| Paperware | 14651 | 4754 | .15 | Plastics (.05) | Cardboard (.02) |
| Plastics | 40837 | 12146 | .12 | Paperware (.02) | Electricity (.01) |
| Refractories | 6020 | 1787 | .11 | Steel (.03) | Cement (.03) |
| Traditional Medicines | 7660 | 2108 | .10 | Pharmaceutical Materials (.07) | Beverages (.01) |
| Valves | 4308 | 1246 | .10 | Valves (.03) | Plastics (.01) |

Table 4 reports some statistics of the products we will investigate. Column (1) reports the number of observations where the output of the product by a firm in a given year is reported in the Output Database. Column (2) lists the number of producers. Column (3) lists the share of observations where information on at least one more product is also reported by the firm in the same year. We term this number as a product's sibling propensity. Columns (4) and (5) report the top two siblings of the products and the shares of producers (out of those listed in Column (2)) that produce each byproduct.

We decide to focus on products in Table 4 since they have a decent number of observations (>4000 in the Output Database) and a low sibling propensity (≤0.15). These industries give us a quite diverse sample. We have both very homogenous products, such as Cement (powder), and products that are quite heterogeneous, such as Traditional Medicines and Furniture (refer to Table 19 for price variations in each industry). Some industries are more concentrated, such as Ferroalloys, while others are populated by many small firms, such as Plastics. We believe that the diverse market structures and production modes of these industries will provide a good environment to test our method.

We use the market value of the output from the Annual Survey of Industrial Firms and the physical quantity of output in the Output Database to calculate the prices. We delete observations with extreme prices to mitigate the problem of misreporting[5] and the probability that the reported product is fundamentally different from the majority of the output in the industry.[6] We deem a price to be extreme if it is below one-twentieth or above twenty times the yearly median price in the industry.

## 6. Results

### 6.1. Technical Details Concerning Estimation

It is a little difficult to estimate a 9-parameter model with the proxy-variable approach. We would like to make sure that our conclusions are not based on faulty optimization results, so we adopt a very elaborate estimation procedure. When we estimate each of the eight models (Equations (12), (13), (30), (31), (32), (33) , (34), and (35)), we try at least 20 batches of initial values for the optimization procedure. For each initial values, we try at least 6 optimization techniques (permutations of Modified Newton-Raphson, Davidon-Fletcher-Powell, Broyden-Fletcher-Goldfarb-Shanno). Among all the 112 estimations we run (8 models by 13 industries plus pooled sample), 106 of them converge to the same results from at least two sets of initial values, and 101 of them converge to the same results from at least five sets of initial values.

In Equations (12) and (13), the inferred markup enter the equations as a logarithm. Some candidate parameters can lead to a negative markup and cause problems to the estimation procedure. To circumvent this problem, we use the second-order Taylor series around 1 to approximate the logarithm of the inferred markup to speed up the estimation. We found that all the estimation results give positive markups for all observations in our sample.

[5] For observations with abnormal prices, we find a high propensity of discrepancies between the major products reported in the Performance Database and that reported in the Output Database.

[6] For example, in the Leather Shoes industry, producers of static dissipative shoes report a much higher average price. In the Plastics industry, producers of medicine capsules also tend to feature much higher average prices.

To lessen the computational burden of our estimation procedure, we normalize monetary values (e.g., revenue, capital, and intermediate inputs), physical quantities of output, and employment with their respective industry-level normalizers.[7]

### 6.2. A Quick Look at Estimated Coefficients

In Table 5 through Table 8, we report the estimated coefficients of various versions of the translog productions functions. The caption of each table indicates the corresponding econometric model. In the proxy-variable approach, optimization is conducted on the fitted values generated in the first stage, the standard errors of the parameters need to be generated with time-consuming bootstrapping.[8] To conserve our computational resources, we do not bootstrap the standard errors. Incidentally, the samples sizes are smaller than those reported in Table 4, since observations need to have additional information from the Performance Database and lagged values for all variables.

Table 5 reports the estimated coefficients when the physical quantity is used as the output measure in a translog production function. We argued at the beginning of this paper that using physical quantity as the output measure without remedies would cause an omitted-variable problem due to the unobservable quality. A quick look at Table 5 will probably lead most of us to suspect that something is indeed wrong. Unlike the usual figures reported for translog estimates (e.g., Berndt and Christensen, 1973, Kim, 1992), the estimated coefficients on $k$ are a large negative number in many cases, those on $l$ are surprisingly large while those on $m$ are unusually small in many cases. The coefficients on second-order terms seem too large in absolute value. The last row reports the results for the pooled sample, which do not show any signs of consistence with results at the disaggregated level. In summary, when the physical quantity is used as the output measure for the translog production function, the estimated parameters seem gravely suspicious.

[7] The product price statistics reported in this paper is based on data that do not go through this procedure.

[8] It takes a single-core processor about 2800 hours to finish a single run of all the estimations.

**Table 5: Estimated Parameters of Production Function: Physical Quantity as Dependent Variable (Equation (32))**

| Product | (1) Obs | (2) $k$ | (3) $l$ | (4) $m$ | (5) $k^2$ | (6) $l^2$ | (7) $m^2$ | (8) $kl$ | (9) $km$ | (10) $lm$ | (11) p-value of Over-identification Test | (12) J Statistic |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Bearings | 2050 | .005 | -.280 | .882 | .040 | -.363 | -.140 | -.079 | -.016 | .530 | .000 | 621.567 |
| Cement | 18574 | .016 | -.892 | 1.095 | -.000 | -.095 | .090 | -.082 | .108 | -.206 | .000 | 4022.078 |
| Ceramics | 2167 | -.137 | .432 | .689 | -.032 | -.095 | -.120 | .007 | .004 | .217 | .000 | 194.517 |
| Engineered Wood | 5681 | .091 | -.135 | .961 | .046 | -.101 | -.057 | -.117 | .043 | .126 | .000 | 199.495 |
| Ferroalloys | 2445 | .017 | .312 | .628 | .006 | .128 | -.063 | -.148 | .100 | -.023 | .000 | 89.663 |
| Furniture | 4772 | -.177 | .993 | .331 | -.035 | .009 | .044 | .071 | .062 | -.176 | .000 | 200.591 |
| Garments | 32164 | -.199 | .393 | .774 | -.032 | .063 | -.060 | .019 | .040 | -.041 | .000 | 388.665 |
| Leather Shoes | 5323 | -.147 | -.554 | 1.274 | -.018 | -.806 | -.593 | -.152 | .136 | 1.364 | .000 | 878.813 |
| Paperware | 6852 | .020 | .278 | .682 | .000 | .247 | .144 | -.062 | .082 | -.471 | .000 | 238.311 |
| Plastics | 19684 | -.086 | -.066 | 1.008 | -.015 | .004 | -.030 | -.039 | .079 | -.008 | .000 | 610.603 |
| Refractories | 3002 | -.084 | .255 | .739 | -.034 | .145 | -.003 | .014 | .042 | -.122 | .000 | 47.191 |
| Traditional Medicines | 3306 | -.126 | .737 | .343 | -.046 | .207 | .167 | .166 | -.041 | -.390 | .000 | 366.114 |
| Valves | 2136 | .147 | -.106 | 1.028 | .030 | -.110 | -.107 | .027 | -.041 | .141 | .000 | 131.549 |
| All Industries | 108156 | 1.433 | -.464 | -.132 | .155 | -.010 | .218 | .142 | -.268 | -.303 | .000 | 116761 |

Column (11) reports the p-values of the overidentification tests. The null is rejected in all cases, suggesting that the orthogonality assumptions are unlikely to be true. We experiment with different sets of moment conditions, but have little success. We turn to the fixed-effect estimator, which is based on the assumption that the unobservable productivity is firm-specific and largely time-invariant. We also try the Arellano-Bond estimator, which estimate a first-difference equation using levels of lagged inputs as instruments. However, both estimators generate equally erratic estimated parameters, returns to scale, and markups.

We also try to estimate the above equation on firms that have always reported a single product, in the hope of mitigating the problem of omission of byproducts. Unfortunately, the results are equally questionable.

Table 6 reports the estimated coefficients when revenue is used as the output measure in a translog production function. Most of the problems observed in Table 5 have vanished. A few red flags remain though. The over-identification tests reject the null for Cement, Plastics, and the pooled sample at the usual significance level, and almost reject it for Refractories.

**Table 6: Estimated Parameters of Production Function: Revenue as Dependent Variable (Equation (33))**

| Product | (1) Obs | (2) $k$ | (3) $l$ | (4) $m$ | (5) $k^2$ | (6) $l^2$ | (7) $m^2$ | (8) $kl$ | (9) $km$ | (10) $lm$ | (11) p-value of Over-identification Test | (12) J Statistic |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Bearings | 2050 | .020 | .003 | .928 | .003 | -.011 | .042 | -.009 | .013 | -.032 | .797 | 3.091 |
| Cement | 18574 | -.001 | .007 | .988 | -.003 | -.006 | -.008 | .010 | .006 | .005 | .000 | 58.672 |
| Ceramics | 2167 | -.005 | .096 | .908 | -.005 | .017 | .028 | .012 | .010 | -.056 | .441 | 5.844 |
| Engineered Wood | 5681 | -.009 | .039 | .968 | -.006 | .006 | .004 | .012 | .006 | -.021 | .195 | 8.633 |
| Ferroalloys | 2445 | .007 | .032 | .939 | -.003 | .007 | .018 | .002 | .002 | -.028 | .999 | .390 |
| Furniture | 4772 | .006 | -.026 | 1.018 | -.003 | -.008 | -.019 | .001 | .001 | .027 | .901 | 2.198 |
| Garments | 32164 | .027 | .104 | .883 | .003 | .028 | .026 | .001 | -.006 | -.051 | .759 | 3.388 |
| Leather Shoes | 5323 | .007 | .115 | .895 | -.000 | .035 | .029 | .004 | .000 | -.067 | .992 | .812 |
| Paperware | 6852 | .014 | .031 | .954 | .001 | .018 | .011 | -.004 | -.001 | -.020 | .922 | 1.980 |
| Plastics | 19684 | .018 | .044 | .944 | .001 | .022 | .010 | -.000 | -.003 | -.027 | .002 | 21.265 |
| Refractories | 3002 | -.033 | .031 | 1.031 | -.011 | .004 | -.020 | .004 | .028 | -.007 | .102 | 10.600 |
| Traditional Medicines | 3306 | .041 | .029 | .946 | .003 | -.006 | .008 | .012 | -.016 | .001 | .608 | 4.511 |
| Valves | 2136 | .040 | .042 | .948 | .009 | .017 | .010 | -.010 | -.001 | -.031 | .942 | 1.738 |
| All Industries | 108156 | .008 | .052 | .947 | -.001 | .013 | .007 | -.001 | .003 | -.021 | .000 | 478.099 |

Table 7 reports the estimated parameters of the translog production function where revenue is used as the output measure and the inferred markup is used as the control function of the unobservable prices and quality. Compared with Table 6, the most obvious and systematic change is that the J statistics, i.e., the objective function of the GMM estimator, decreases significantly, indicating an increase in the goodness of fit. As a result, none of the over-identification tests reject the null at the usual significance level.

The underlying econometric model of Table 8 is very similar to that of Table 7. The only difference is that in the new model restricts that variations in the flexible input has no effect on quality, i.e., $\frac{\partial \ln Q_{it}}{\partial m_{it}} = 0$. The change speeds up estimation about 40 percent,[9] but the gain in estimation speed comes with the cost of goodness of fit. The J statistics increases in almost all industries and causes the over-identification test to reject the null in the Garments industry.

[9] An estimation run for a set of initial values and optimization techniqe based on Equation (13) takes about 475 seconds on average, while that based on Equation Equation (12) takes about 666 seconds.

**Table 7: Estimated Parameters of Production Function: Revenue as Dependent Variable, Markup as Control (Equation (12))**

| Product | (1) Obs | (2) $k$ | (3) $l$ | (4) $m$ | (5) $k^2$ | (6) $l^2$ | (7) $m^2$ | (8) $kl$ | (9) $km$ | (10) $lm$ | (11) p-value of Over-identification Test | (12) J Statistic |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Bearings | 2050 | .000 | .011 | .976 | .001 | .000 | .004 | -.001 | .001 | -.004 | 1.000 | .021 |
| Cement | 18574 | -.002 | .002 | .998 | -.000 | -.000 | -.000 | .001 | .001 | -.000 | .999 | .352 |
| Ceramics | 2167 | -.002 | .025 | .975 | -.001 | .003 | .004 | .002 | .001 | -.007 | 1.000 | .076 |
| Engineered Wood | 5681 | -.002 | .009 | .995 | -.001 | .002 | .001 | .000 | .001 | -.003 | 1.000 | .063 |
| Ferroalloys | 2445 | .000 | .011 | .985 | -.000 | .001 | .002 | .000 | .000 | -.004 | 1.000 | .010 |
| Furniture | 4772 | .000 | .015 | .987 | -.000 | .002 | .002 | .001 | .000 | -.004 | 1.000 | .080 |
| Garments | 32164 | .001 | .048 | .954 | -.000 | .007 | .007 | .000 | .000 | -.014 | .217 | 8.297 |
| Leather Shoes | 5323 | .001 | .086 | .933 | -.000 | .019 | .012 | .002 | -.001 | -.031 | .999 | .422 |
| Paperware | 6852 | .002 | .021 | .979 | .000 | .005 | .004 | .000 | -.000 | -.008 | 1.000 | .032 |
| Plastics | 19684 | .004 | .026 | .970 | .000 | .004 | .005 | .001 | -.001 | -.008 | .991 | .827 |
| Refractories | 3002 | -.001 | .030 | .980 | -.000 | .005 | .002 | -.001 | .001 | -.008 | 1.000 | .069 |
| Traditional Medicines | 3306 | .011 | .020 | .976 | .001 | .004 | .003 | .001 | -.002 | -.006 | .892 | 2.279 |
| Valves | 2136 | .005 | .011 | .989 | .001 | .002 | .001 | -.000 | -.001 | -.003 | 1.000 | .013 |
| All Industries | 108156 | -.000 | .017 | .985 | -.000 | .003 | .002 | -.000 | .000 | -.005 | .341 | 6.787 |

**Table 8: Estimated Parameters of Production Function: Revenue as Dependent Variable, Corrected Markup as Control (Equation (13))**

| Product | (1) Obs | (2) $k$ | (3) $l$ | (4) $m$ | (5) $k^2$ | (6) $l^2$ | (7) $m^2$ | (8) $kl$ | (9) $km$ | (10) $lm$ | (11) p-value of Over-identification Test | (12) J Statistic |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Bearings | 2050 | .001 | .007 | .982 | .001 | .000 | .005 | -.001 | .001 | -.005 | 1.000 | .031 |
| Cement | 18574 | -.001 | .001 | .998 | -.000 | -.000 | -.000 | .001 | .001 | -.000 | .999 | .344 |
| Ceramics | 2167 | -.001 | .020 | .979 | -.001 | .003 | .004 | .002 | .001 | -.008 | 1.000 | .093 |
| Engineered Wood | 5681 | -.002 | .007 | .995 | -.001 | .002 | .001 | .001 | .001 | -.003 | 1.000 | .066 |
| Ferroalloys | 2445 | .001 | .008 | .988 | -.000 | .001 | .002 | .000 | .000 | -.004 | 1.000 | .014 |
| Furniture | 4772 | .003 | -.026 | 1.018 | -.000 | -.010 | -.007 | .000 | -.001 | .017 | 1.000 | .105 |
| Garments | 32164 | .001 | .042 | .960 | -.000 | .009 | .008 | .000 | .000 | -.018 | .020 | 14.998 |
| Leather Shoes | 5323 | .000 | .078 | .941 | -.001 | .028 | .017 | .002 | -.001 | -.045 | .982 | 1.082 |
| Paperware | 6852 | .002 | .016 | .984 | .000 | .006 | .004 | .000 | -.000 | -.009 | 1.000 | .048 |
| Plastics | 19684 | .000 | .005 | .993 | .000 | .002 | .001 | -.001 | .000 | -.002 | .974 | 1.251 |
| Refractories | 3002 | -.003 | .003 | 1.002 | -.001 | .000 | -.001 | .000 | .002 | -.000 | 1.000 | .068 |
| Traditional Medicines | 3306 | .010 | .015 | .980 | .001 | .004 | .004 | .001 | -.003 | -.006 | .810 | 2.988 |
| Valves | 2136 | .005 | .008 | .992 | .001 | .002 | .001 | -.000 | -.001 | -.003 | 1.000 | .014 |
| All Industries | 108156 | .000 | .014 | .987 | -.000 | .003 | .002 | -.000 | .000 | -.006 | .193 | 8.674 |

To quantify the differences between the estimated parameters in Table 5 through Table 8, we test the hypothesis that a parameter estimated by two methods are the same. Since we do not have the bootstrap stand errors, we use the GMM standard errors instead. The results will be inaccurate, but will give us some perspective about the differences between the results of different models.

Table 9 reports the number of rejections out of 126 tests (9 parameters by 14 samples) against the null that parameters in a pair of models are identical. The significance level is 5 percent. Column (1) shows that the null of parameter equality is rejected about 70 percent of the time when the results based on physical quantity are compared with others. The Column (2) shows that adding a control function to the revenue production function makes significant differences. The null of parameter equality is reject about 65 percent of the time. The very last cell of the table shows that the null is rejected about 23 percent of the time when the estimates of two versions of our methods are compared.

**Table 9: Number of Rejected Nulls of Equality in Parameters at 5-Percent Significance Level**

| | (1)<br>• Quantity as Output<br>• Translog<br>• Equation (32) | (2)<br>• Revenue as Output<br>• Translog<br>• Equation (33) | (3)<br>• Revenue as Output<br>• Translog+log(markup)<br>• Equation (12) |
|---|---|---|---|
| • Revenue as Output<br>• Translog<br>• Equation (33) | 89 | | |
| • Revenue as Output<br>• Translog+log(markup)<br>• Equation (12) | 92 | 81 | |
| • Revenue as Output<br>• Translog + log(corrected markup)<br>• Equation (13) | 90 | 83 | 29 |

## 6.3. Estimated Returns-to-Scale and Markups

In this section we report statistics of inferred returns-to-scale and markups of various models. We intend to demonstrate (1) that the inferred returns-scale and the markups are very different from our expectation when the physical quantity is used as output measure without remedies for unobservable quality in the estimation of production functions, (2) that the two inferred measures changes considerably when we add a control function for unobservable prices and qualities in the production function with revenue as the output measure.

#### 6.3.1. Estimated Returns to Scale

Table 10 reports statistics of the estimated returns-to-scale for various models. To save space, we report four statistics in each cell. The first number in each cell is the median of the estimated returns-to-scale; the first number inside the parentheses is the standard deviation of the estimated returns-to-scale; the next two numbers are minimum and the maximum of the estimated returns-to-scale. In almost all cases, the median is indistinguishable from the mean (not reported here). Columns (1) and (2) contain information that has being reported in Table 3, but provide more details.

**Table 10: Statistics of Estimated Returns to Scale**

| Product | (1)<br>• Quantity as Output<br>• Translog<br>• Equation (32)<br>Median(SE,Min,Max) | (2)<br>• Revenue as Output<br>• Translog<br>• Equation (33)<br>Median(SE,Min,Max) | (3)<br>• Revenue as Output<br>• Translog+log(markup)<br>• Equation (12)<br>Median(SE,Min,Max) | (4)<br>• Revenue as Output<br>• Translog + log(corrected markup)<br>• Equation (13)<br>Median(SE,Min,Max) |
|---|---|---|---|---|
| Bearings | .71 (.22,-.07,1.53) | .97 (.05,.76,1.20) | .99 (.00,.97,1.01) | 1.00 (.00,1.00,1.00) |
| Cement | .24 (.32,-1.08,1.77) | .99 (.01,.93,1.05) | 1.00 (.00,.99,1.01) | 1.00 (.00,1.00,1.00) |
| Ceramics | 1.02 (.06,.74,1.26) | .99 (.02,.93,1.07) | 1.00 (.00,.99,1.01) | 1.00 (.00,.99,1.01) |
| Engineered Wood | .94 (.12,.36,1.44) | .99 (.01,.97,1.02) | 1.00 (.00,1.00,1.00) | 1.00 (.00,1.00,1.00) |
| Ferroalloys | .96 (.06,.69,1.28) | .98 (.01,.93,1.02) | 1.00 (.00,.99,1.00) | 1.00 (.00,1.00,1.00) |
| Furniture | 1.08 (.09,.75,1.49) | 1.00 (.01,.92,1.04) | 1.00 (.00,1.00,1.00) | 1.00 (.00,1.00,1.00) |
| Garments | .90 (.10,.10,1.39) | 1.01 (.00,.98,1.03) | 1.00 (.00,1.00,1.00) | 1.01 (.00,.99,1.02) |
| Leather Shoes | .84 (.31,-.36,3.41) | 1.01 (.01,.96,1.04) | 1.02 (.01,.95,1.04) | 1.01 (.00,.99,1.01) |
| Paperware | .92 (.12,.35,1.55) | 1.00 (.01,.95,1.04) | 1.00 (.00,.99,1.01) | 1.00 (.00,1.00,1.00) |
| Plastics | .86 (.03,.72,1.02) | 1.00 (.01,.93,1.06) | 1.00 (.00,1.00,1.00) | 1.00 (.00,.99,1.00) |
| Refractories | .87 (.14,.42,1.35) | 1.01 (.02,.96,1.10) | 1.01 (.00,1.00,1.02) | 1.00 (.00,.99,1.00) |
| Traditional Medicines | .93 (.14,.31,1.41) | 1.02 (.00,.99,1.03) | 1.01 (.00,.99,1.01) | 1.00 (.00,.99,1.01) |
| Valves | .96 (.12,.48,1.43) | 1.02 (.01,.97,1.07) | 1.00 (.00,1.00,1.01) | 1.00 (.00,1.00,1.00) |
| All Industries | .63 (.24,-.68,1.94) | 1.01 (.00,.98,1.03) | 1.00 (.00,1.00,1.00) | 1.00 (.00,1.00,1.00) |

As it was mentioned in Section 2.2, when the physical quantity is used as the output measure without remedies for unobservable quality (Column (1)), the estimated returns to scale are quite different from unit in many industries. There are also substantial variations in returns-to-scale across firms, which is evidenced by the standard error and the range. In Bearings, Cement, Leather Shoes and the pooled sample, some firms have negative estimated returns-to-scale, which cannot be rationalized. In other samples, some firms have extremely low or high estimated returns-to-scale. These extreme values of estimated returns-to-scale are hard to justify, since they indicate that firms should find it highly profitable to adjust their scale. We believe the erratic estimated returns-to-scale are an acute symptom of the omitted-variable problem caused by the unobservable quality.

Column (2) shows the statistics of the returns-to-scale when revenue is used as the output measure. The medians of the estimated returns-to-scale are very close to unit and none of the firms have returns-to-scale that are far from the median.

Columns (3) and (4) show the estimated returns-to-scale based on the models that we propose, i.e., using the logarithm of the markup as the control function of unobservable prices and quality when the revenue is the available output measure. The estimated returns-to-scale appear to be reasonable. Compared with results in Column (2), estimated returns-to-scale generated by our models are more concentrated around unit.

Results in Columns (2) – (4) suggest that almost all firms exhibit strict constant returns to scale. In a world without adjustment costs, this implies the marginal cost is constant and is equal to the average costs. This will lend more support to our assumption that quality is a function of the marginal cost and productivity. It will also render the markup to be exactly equal to the price margin, making it appropriate to view revenue as a function of inputs, productivity and markup in theory.

**Table 11: Correlation between Returns-to-Scale Based on Various Models**

| | •Quantity as Output<br>•Translog<br>•Equation (32) | •Revenue as Output<br>•Translog<br>•Equation (33) | •Revenue as Output<br>•Translog+log(markup)<br>•Equation (12) |
|---|---|---|---|
| •Revenue as Output<br>•Translog<br>•Equation (33) | -.06 | | |
| •Revenue as Output<br>•Translog+log(markup)<br>•Equation (12) | -.22 | .46 | |
| •Revenue as Output<br>•Translog +<br>log(corrected markup)<br>•Equation (13) | .04 | .52 | .70 |

Table 11 reports the correlation of firms' estimated returns-to-scale based on estimates of various models at the industry level. We can see that the estimated returns-to-scale based Equation (32) (quantity as output) appear to be at odds with those based on other models. The estimated returns to scale based on Equation (33) (Revenue as output, no further corrections) are highly correlated with those generated by our models, but far from identical. Finally, estimated returns to scale generated our two models seem to be consistent. We do not want to read too much into the correlation coefficients involving our models too much though, since the estimated returns to scale generated by them are almost exactly equal to unit for most firms. The lack of variation makes the correlation coefficient lose its economic significance.

Table 12 reports some statistics of the estimated returns-to-scale generated by disaggregated and pooled samples for various models. When the physical quantity is used as the output measure, the mean returns to scale is below unit and the dispersion is large. For the next three models reported in the table, the mean progressively converges to one and the standard error successively decreases. In the last row, we report the correlation of the estimated returns to scale generated by the disaggregate samples and the pooled sample. Our models (Columns (3) and (4)) report a slightly lower coefficient of correlation. This is probably because there is not much variation in the returns to scale estimated by our models (refer to Table 10 for more details on standard deviations).

**Table 12: Statistics of Returns to Scale Generated by Disaggregated and Pooled Samples**

| | (1)<br>• Quantity as Output<br>• Translog<br>• Equation (32) | (2)<br>• Revenue as Output<br>• Translog<br>• Equation (33) | (3)<br>• Revenue as Output<br>• Translog + log(markup)<br>• Equation (12) | (4)<br>• Revenue as Output<br>• Translog + log (corrected markup)<br>• Equation (13) |
|---|---|---|---|---|
| Mean based on industry-level estimation | .942 | 1.003 | 1.001 | 1.001 |
| Mean based on pooled sample estimation | .627 | 1.006 | 1.000 | 1.001 |
| Standard error based on industry-level estimation | .190 | .013 | .007 | .005 |
| Standard error based on pooled sample estimation | .195 | .005 | .001 | .001 |
| Correlation between returns to scale based on two estimations | .304 | .369 | .225 | .264 |

### 6.3.2. Estimated Markups

Table 13 reports statistics of the estimated markups for various models. The structure of the table is the same as that of Table 10, except that we also report the percentage of observations with a markup greater than one.

Results in Column (1) give some quite unbelievable statistics on markups when the physical quantity is used as output without remedies for unobservable quality. Many industries have markups that are far below unit. Firms in some industries even have negative markups. Overall, we believe that the symptoms in estimated markups signals the severity of the omitted-variable problem caused by unobservable quality when the physical quantity is used as output.

Results in Column (2) are much more well-behaved. The median markup is in the range of 1.19 to 1.36, which seem reasonable. Most firms have a markup greater than one, but in Bearings , Ceramics, Furniture, and Refractories, there are a tiny but noticeable portion of firms that have a markup below one. .

**Table 13: Statistics of Estimated Markups**

| Product | (1)<br>• Quantity as Output<br>• Translog<br>• Equation (32)<br>Median<br>(SE,Min,Max,%>1) | (2)<br>• Revenue as Output<br>• Translog<br>• Equation (33)<br>Median<br>(SE,Min,Max,%>1) | (3)<br>• Revenue as Output<br>• Translog+log(markup)<br>• Equation (12)<br>Median<br>(SE,Min,Max,%>1) | (4)<br>• Revenue as Output<br>• Translog + log(corrected markup)<br>• Equation (13)<br>Median<br>(SE,Min,Max,%>1) |
|---|---|---|---|---|
| Bearings | 1.05 (.56,-.82,2.90,.54) | 1.30 (.21,.55,2.42,.99) | 1.30 (.06,1,1.54,1) | 1.35 (.06,1.09,1.69,1) |
| Cement | 1.49 (.29,.16,3.37,.97) | 1.27 (.07,.85,1.93,1) | 1.33 (.05,.95,1.89,1) | 1.33 (.05,.95,1.88,1) |
| Ceramics | .84 (.33,-.34,2.27,.31) | 1.24 (.10,.57,1.99,.99) | 1.34 (.06,.78,1.60,1) | 1.37 (.07,.82,1.68,1) |
| Engineered Wood | 1.15 (.17,.50,2.23,.83) | 1.27 (.04,.91,1.53,1) | 1.33 (.04,.88,1.68,1) | 1.33 (.04,.88,1.69,1) |
| Ferroalloys | .63 (.19,-.14,1.52,.01) | 1.28 (.07,.80,1.64,1) | 1.32 (.04,.99,1.60,1) | 1.35 (.04,1,1.66,1) |
| Furniture | .43 (.16,-.40,1.52,.00) | 1.25 (.15,.64,2.33,.99) | 1.29 (.06,.90,1.87,1) | 1.31 (.06,.93,1.92,1) |
| Garments | .86 (.21,-.13,3.18,.24) | 1.19 (.06,.85,1.83,1) | 1.24 (.07,1,2.61,1) | 1.18 (.05,1.01,2.04,1) |
| Leather Shoes | .44 (1.52,-1.62,1.02,.33) | 1.22 (.07,.55,2.21,1) | 1.23 (.05,.69,1.89,1) | 1.23 (.07,.73,2.10,1) |
| Paperware | 1.01 (.40,-1.33,3.38,.51) | 1.23 (.03,.68,1.53,1) | 1.24 (.04,.77,1.68,1) | 1.26 (.05,.80,1.77,1) |
| Plastics | 1.17 (.14,.63,1.91,.91) | 1.20 (.03,.78,1.82,1) | 1.23 (.04,.88,1.95,1) | 1.24 (.04,.88,2.00,1) |
| Refractories | .96 (.13,.39,1.52,.38) | 1.28 (.14,.77,2.12,.99) | 1.32 (.04,.74,1.57,1) | 1.36 (.05,.75,1.66,1) |
| Traditional Medicines | .60 (.42,-1.56,2.52,.18) | 1.36 (.06,.89,2.57,1) | 1.42 (.07,.97,2.56,1) | 1.51 (.08,1,2.68,1) |
| Valves | 1.22 (.30,.22,2.80,.76) | 1.22 (.05,.89,1.67,1) | 1.27 (.06,1.02,1.72,1) | 1.28 (.06,1.02,1.75,1) |
| All Industries | .41 (.60,-3.17,2.56,.15) | 1.22 (.04,.75,1.99,1) | 1.29 (.06,.80,2.42,1) | 1.29 (.06,.80,2.42,1) |

Results in Columns (3) and (4) are based on our models that uses the logarithm of markup as the control function of unobservable prices and quality. The markups based on these models are similar to those based the model underlying Column (2), but differ from the latter in three respects. First, the markups generated by our models are on average greater. Second, the markups generated by our models are less dispersed. Finally, almost all firms have an estimated markup greater than one.

Table 14 reports the correlation of the estimated markups based on estimates of various models at the industry level. The markups generated by Equation (32) (quantity as output) only has very weak correlation with those generated by other models. The estimated markups based on Equation (33) (revenue as output, no further corrections) are moderately correlated with those generated by our models. Finally, the large number in the very last cell indicates the internal consistency of our two models.

**Table 14: Correlation between Markups Based on Various Models**

| | •Quantity as Output<br>•Translog<br>•Equation (32) | •Revenue as Output<br>•Translog<br>•Equation (33) | •Revenue as Output<br>•Translog+log(markup)<br>•Equation (12) |
|---|---|---|---|
| •Revenue as Output<br>•Translog<br>•Equation (33) | .07 | | |
| •Revenue as Output<br>•Translog+log(markup)<br>•Equation (12) | .20 | .33 | |
| •Revenue as Output<br>•Translog + log(corrected markup)<br>•Equation (13) | .06 | .34 | .92 |

Since economists use data at various aggregation level, they would like their econometric model to be robust to the aggregation. We checked the correlation of estimated returns to scale generated at different aggregation level, but the results are not very informative since the variation in the estimated returns to scale is extremely small. Table 13 shows that there is much more variations in the estimated markups than those reported in Table 10 for estimated returns to scale. In what follows, we check the correlation of estimated markups at different aggregation level and hope it can shed light on the validity of our models.

**Table 15: Correlation of Markups Generated by Disaggregated and Pooled Samples**

| | (1)<br>• Quantity as Output<br>• Translog<br>• Equation (32) | (2)<br>• Revenue as Output<br>• Translog<br>• Equation (33) | (3)<br>• Revenue as Output<br>• Translog + log(markup)<br>• Equation (12) | (4)<br>• Revenue as Output<br>• Translog + log (corrected markup)<br>• Equation (13) |
|---|---|---|---|---|
| Mean based on industry-level estimation | 1.037 | 1.226 | 1.265 | 1.301 |
| Mean based on pooled sample estimation | .591 | 1.207 | 1.297 | 1.315 |
| Standard error based on industry-level estimation | .489 | .098 | .070 | .074 |
| Standard error based on pooled sample estimation | .448 | .065 | .064 | .068 |
| Mean based on industry-level estimation | -.293 | .219 | .715 | .829 |

Table 15 reports some statistics of the estimated markups generated by disaggregated and pooled samples for various models. We can see that using physical quantity as output without remedies for unobservable quality lead to a questionable negative correlation between the markups generated by disaggregated and pooled samples for the same firm. The mercuriality of the markups is probably due to the changes of the underlying quality metric in different industries, to which the econometric model is likely to be susceptible. Column (2) shows that when we use revenue as output without further corrections, there is a moderate correlation between the markups generated at different aggregation levels, indicating that estimates are probably more robust to the aggregation level. Columns (3) and (4) shows that our models are much more robust to the changes in the aggregation level, viewed through the lens of the estimated markup.

In summary, the estimated returns-to-scale and markups are quite questionable when the physical output is used output without further corrections. Using revenue as output would greatly alleviate the symptoms. Adding the logarithm of markup as the control function of prices and quality would generate correlated but appreciably different returns-to-scale and markups, suggesting the importance of using our models. Among the models discussed above, our approach give much higher correlation between markups generated by disaggregated and pooled samples for the sample, showing that our models is more robust to the changes in the quality normalizer in different industries.

### 6.4. Coefficient on Control Function

The difference between our approach and the traditional approach can be viewed as a matter of whether the coefficient on the control function is one or zero. Our approach holds that the coefficient on the control function, $\beta_{Markup}$, is equal to one (see Equation (4), or Equations (12) and (13)), while the traditional approach implicitly assume that $\beta_{Markup}$ is equal to zero (see Equation (33)). In this section, we check which theory has more support from the data.

We estimate Equations (34) and (35), where the coefficient on the control function, $\ln \mu$, is treated as a free parameter. We use the estimated coefficients for Equations (12) and (13), and ten arbitrary values for $\beta_{Markup}$, as the initial values for the estimation procedure. In Table 16, we report the estimated value of $\beta_{Markup}$. In parentheses, we report the asymptotic standard error of the GMM estimator. The asterisks after the parentheses indicate the significance level of the test of the null $\beta_{Markup} = 1$ using the asymptotic standard errors.

**Table 16: Estimated Coefficient on Markup in Equation (34) and (35)**

| Product | (1)<br>$\beta_{Markup}$ in Equation (34) | (2)<br>$\beta_{Markup}$ in Equation (35) (Corrected Markup) |
|---|---|---|
| Bearings | 1.09 (.02)*** | 1.09 (.01)*** |
| Cement | 1.11 (.02)*** | 1.11 (.02)*** |
| Ceramics | 1.18 (.05)*** | 1.19 (.05)*** |
| Engineered Wood | 1.11 (.04)*** | 1.12 (.04)*** |
| Ferroalloys | 1.10 (.02)*** | 1.10 (.02)*** |
| Furniture | 1.19 (.05)*** | 1.21 (.05)*** |
| Garments | .48 (.23)*** | .22 (.33)*** |
| Leather Shoes | .86 (.14)* | .81 (.18)* |
| Paperware | 1.10 (.03)*** | 1.11 (.03)*** |
| Plastics | 1.20 (.08)*** | 1.25 (.10)*** |
| Refractories | 1.10 (.01)*** | 1.10 (.01)*** |
| Traditional Medicines | 2.24 (.69)** | 2.43 (.81)** |
| Valves | 1.12 (.04)*** | 1.12 (.03)*** |
| all Industries | .99 (.05) | 1.00 (.06) |

- The asterisks following the parentheses indicate the significance level of the test $\beta_{Markup} = 1$.
- ***, **, and * indicates significance levels of 1%, 5%, and 10% respectively

Our first observation from Table 16 is that the estimated values of $\beta_{Markup}$ are obviously closer to the value predicted by our theory, namely, one, than the one implicitly assumed by the traditional approach, that is, zero. If we need to decide whether to include $\ln \mu$ in the production function when revenue is used as output, then the results in Table 16 suggest that the answer is definitely in the affirmative.

When we use the asymptotic GMM standard errors to test the hypothesis $\beta_{Markup} = 1$, we reject the hypothesis at the 5% level in 11 out of 22 cases. We are unfazed by the high rejection rate though, since in quite a few cases the null is rejected because the standard errors are rather small. In these cases, the small standard errors, in conjunction with the estimated values of $\beta_{Markup}$, assure us that the population coefficient of control function is close to one.[10]

Overall, we take results in Table 16 as encouraging news for the method that we propose. It shows that our approach is probably more appropriate than the traditional approach. Moreover, it assures us that our approach is not merely a mediocre remedy of the traditional approach, but probably a rather satisfactory treatment of the problem posed by the unobserved prices and qualities, since $\hat{\beta}_{Markup}$ is close to the theoretic value.

### 6.5. Model and Moment Selection Criteria

Andrews and Lu (2001) develop a set of consistent Model and Moment Selection Criteria (MMSC) for GMM estimation. Their criteria are based on the J statistic of GMM estimations and can select the correct model specification asymptotically. These criteria are similar to likelihood-based selection criteria AIC and BIC.

In this work, we will use Andrews and Lu's MMSC-AIC and MMSC-BIC to compare various models. Like traditional information criteria AIC and BIC, a model with a lower MMSC-AIC or MMSC-BIC is preferred. For details on how to compute the two criteria, we refer readers to Andrews and Lu (2001).

[10] For the Traditional Medicines industry, $\hat{\beta}_{Markup}$ is quite large, but it also appears that there must be something special about this industry, since the standard error of $\hat{\beta}_{Markup}$ is considerably greater than in other industries.

Table 17 reports the MMSC-AIC of models that use revenue as the dependent variable. In the top row, we give a brief hint of the model specifications and the indexes of equations underlying the models. Judging from the MMSC-AIC scores, among the models that use revenue as the output measure, the Cobb-Douglas production function (Column (1)) is least preferred. It has by far the highest MMSC-AIC score for all industries. Adding a control function with its coefficient set to one (Columns (3) and (4)) leads to a lower MMSC-AIC than a traditional translog function does (Column (2)) for almost all industries. Correcting the markup does not cause significant changes in the information score, since the figures in Columns (3) and (4) are quite close. Columns (5) and (6) reports the MMSC-AIC where the coefficient on the control function is set as a free parameter. The MMSC-AIC in these two columns are greater than those in Columns (3) and (4) for almost all industries, showing that the gain in the goodness of fit is outweighed by the loss of simplicity of the model (adding a parameter). Indeed, results in Table 16 show that the estimated coefficient on the control function is quite close to the value preset in our models (Columns (3) and (4)), so there is little to gain by setting it as a free parameter.

**Table 17: MMSC-AIC of Models Using Revenue as Output Measure**

| Product | (1) Cobb-Douglas<br>Equation (31) | (2) Translog<br>Equation (33) | (3) Translog + log(markup)<br>Equation (12) | (4) Translog + log(corrected markup)<br>Equation (13) | (5) Translog + $\beta$*log(markup)<br>Equation (34) | (6) Translog + $\beta$*log(corrected markup)<br>Equation (35) |
|---|---|---|---|---|---|---|
| Bearings | 57.82 | -8.91 | -11.98 | -11.97 | -10.00 | -10.00 |
| Cement | 1535.47 | 46.67 | -11.65 | -11.66 | -9.98 | -9.98 |
| Ceramics | -2.65 | -6.16 | -11.92 | -11.91 | -10.00 | -10.00 |
| Engineered Wood | 276.10 | -3.37 | -11.94 | -11.93 | -10.00 | -10.00 |
| Ferroalloys | -.38 | -11.61 | -11.99 | -11.99 | -10.00 | -10.00 |
| Furniture | 411.72 | -9.80 | -11.92 | -11.90 | -9.98 | -9.98 |
| Garments | 1662.20 | -8.61 | -3.70 | 3.00 | -8.31 | -7.09 |
| Leather Shoes | 38.88 | -11.19 | -11.58 | -10.92 | -9.96 | -9.93 |
| Paperware | 226.14 | -10.02 | -11.97 | -11.95 | -9.99 | -9.99 |
| Plastics | 230.72 | 9.26 | -11.17 | -10.75 | -9.51 | -9.27 |
| Refractories | 1522.82 | -1.40 | -11.93 | -11.93 | -9.99 | -9.99 |
| Traditional Medicines | 25.02 | -7.49 | -9.72 | -9.01 | -9.64 | -9.49 |
| Valves | 58.03 | -10.26 | -11.99 | -11.99 | -10.00 | -10.00 |
| All Industries | 4517.25 | 466.10 | -5.21 | -3.33 | -3.23 | -1.33 |

Table 18 reports the MMSC-BIC of models that have revenue as the dependent variable. The patterns are very similar to those observed in Table 17 except that the Cobb-Douglas function fares very well in Ceramics, Ferroalloys, and Traditional Medicines.

Table 18: MMSC-BIC of Models Using Revenue as Output Measure

| Product | (1) Cobb-Douglas<br>Equation (31) | (2) Translog<br>Equation (33) | (3) Translog + log(markup)<br>Equation (12) | (4) Translog + log(corrected markup)<br>Equation (13) | (5) Translog + $\beta$*log(markup)<br>Equation (34) | (6) Translog + $\beta$*log(corrected markup)<br>Equation (35) |
|---|---|---|---|---|---|---|
| Bearings | -9.68 | -42.66 | -45.73 | -45.72 | -38.13 | -38.13 |
| Cement | 1441.51 | -0.30 | -58.63 | -58.63 | -49.13 | -49.13 |
| Ceramics | -70.83 | -40.24 | -46.01 | -45.99 | -38.41 | -38.41 |
| Engineered Wood | 196.36 | -43.24 | -51.81 | -51.80 | -43.22 | -43.22 |
| Ferroalloys | -70.00 | -46.42 | -46.80 | -46.80 | -39.01 | -39.01 |
| Furniture | 334.07 | -48.62 | -50.74 | -50.72 | -42.33 | -42.33 |
| Garments | 1561.66 | -58.88 | -53.97 | -47.27 | -50.20 | -48.98 |
| Leather Shoes | -40.08 | -50.67 | -51.06 | -50.40 | -42.86 | -42.83 |
| Paperware | 144.15 | -51.01 | -52.96 | -52.95 | -44.15 | -44.15 |
| Plastics | 136.07 | -38.06 | -58.50 | -58.07 | -48.95 | -48.71 |
| Refractories | 1450.74 | -37.44 | -47.97 | -47.97 | -40.02 | -40.02 |
| Traditional Medicines | -48.22 | -44.11 | -46.34 | -45.63 | -40.15 | -40.01 |
| Valves | -9.97 | -44.26 | -45.99 | -45.99 | -38.33 | -38.33 |
| All Industries | 4402.15 | 408.55 | -62.76 | -60.87 | -51.18 | -49.28 |

To summarize, based on the MMSC-AIC and the MMSC-BIC, the models that we propose (Equation (12) or (13)) are most preferred among all the production function models that have revenue as the dependent variable.

## 7. Discussion

In this section, we discuss a few questions that are related to our work but do not seem to fall in a particular section elsewhere in this paper.

### 7.1. Unobserved Quality: Misnomer of an Equally Difficult Problem?

We would like to discuss an alternative explanation for the symptoms of the estimations when the physical quantity is used as the output measure in the production function.

Our analysis on the symptoms is based on the assumption that the physical quantity of the output does not involve severe measurement error. As Roberts and Supina (2000) find out, however, virtually all plants produce more than one output in practice. It is likely that the single-product firms that we focus on actually produce some amounts of byproducts that are not included in the Output Database. If the omission of byproducts in the data varies systematically with some variables in the production function, it will cause bias in the estimates.

As a robustness test, we estimate Equation (32) with data on firms that always report information on a single product, hoping that this is a strong signal that the firm does not have any other product worth reporting. Consequently, the sample size shrinks by about 4 percent. The estimation results are equally erratic as those in Table 3, Table 5, Table 10, and Table 13 for the production function that uses physical quantities as the output measure.

If the robustness test does not dissipate our suspicion that the ill-behaved results may be due to unreported byproducts, then we are led to an inescapable dilemma when the physical quantity is used as the output measure, since we need to choose to deal with unobservable quality or to deal with the missing data on byproducts. Data on quality is seldom available, and the all-inclusive data on the quantities of all products of firms are just as scarce. Even if the all-inclusive output quantity data exist, converting them into a single output measure is probably just as difficult as converting physical output into quality-adjusted output, if not more difficult.

Therefore, it might be a misnomer to refer to the omitted variable as product quality when the physical quantity is used as the output measure, but the underlying measurement problems associated with the physical quantity of output are real and equally difficult to solve.

### 7.2. Recovering Productivity

When we make the assumption that firm $i$'s product quality $Q_{it}$ equals $MC_{it} \cdot \exp\left(\zeta(\omega_{it} + \epsilon_{it})\right)$, we introduce a new parameter $\zeta$. Our method does not provide a way to identity $\zeta$, so the logarithm of productivity is only identifiable up to scale. That is, we can only identify $(1-\zeta)\omega_{it}$. We have two remarks regarding the productivity in our model.

First of all, with the problems of unobservable prices and quality being addressed, now we at least have more confidence in our estimates of $(1-\zeta)\omega_{it}$. In the traditional approach where unobservable prices and quality are left discussed, the recovered productivity is most likely contaminated.

Second, the recovered productivity increases with the firm size, suggesting that it is an increasing function of the true productivity. We regress the estimated value of $(1-\zeta)\omega_{it}$ on the logarithm of the capital stock ($k$) and find that the coefficient on $k$ are positive in 8 out of 10 industries and positive in the pooled sample. The estimated coefficient is economically significant, falling in the range of 0.01-0.02. It is statistically significant in almost all cases.

### 7.3. Constant Quality and Variable Marginal Cost

When we operationalize our assumption on the relation between quality and unit costs, we assume that quality is a function of marginal cost, that is, $Q_{it} = MC_{it} \cdot \exp\big(\zeta(\omega_{it} + \epsilon_{it})\big)$. This leads to a contradiction between constant product quality and variable marginal cost for a given production run. Considering the evidence in our empirical results, we do not think this contradiction a serious empirical problem.

First of all, in most cases the estimated returns to scale are very close to one, so the marginal cost curve in each production run is probably very flat for firms in our sample. If the changes in marginal cost is negligible, the variation in quality loses its importance as a concern.

Second, there is little difference between the results with and without the restriction $\frac{\partial \ln Q_{it}}{\partial m_{it}} = 0$. There are only two systematic differences between two sets of results. The model with the restriction of constant quality seems to fit the data slightly better, judging from the J statistics from Table 7 and Table 8. The model with the restriction tend to be produce more consistent results as the aggregation level changes, evidence by the coefficients of correlation in Table 12 and Table 15.

Third, the value of $\frac{\partial \ln Q_{it}}{\partial m_{it}}$ implied by our model is very small. We compute $\frac{\partial \ln Q_{it}}{\partial m_{it}}$ with the estimated parameters of Equation (12) using Equation (11). The term has a mean of 0.014 and a standard error of 0.012, distributed over the range [-0.175, 0.111]. These figures assure us that the implication of variable quality arising from our assumption is not a serious empirical concern.

### 7.4. Moment Conditions Concerning Markups

In unreported trials, we include in our GMM estimators a few moment conditions involving markups. Most of these moment conditions are Olley-Pakes style orthogonality conditions between innovations in markups and lagged inputs. The additional orthogonal conditions can help the estimator converge much faster. However, in some cases they generate estimates that do not seem "reasonable". It is beyond the scope of this paper to develop theories for the moment conditions concerning markups, so we do not conduct further investigations. Nonetheless, we would like to suggest that this might be worth further analysis.

## 8. Conclusion

In this work, we develop a method to consistently estimate production functions when output prices and product quality are not observed. We show that under some weak assumptions the markup can serve as a control function for unobserved prices and quality. We utilize the theory first developed by Hall (1986, 1988, and 1990) to write the markup as a function of production function parameters. We implement our method as an extension of the proxy-variable approach of estimating production functions pioneered by Olley and Pakes (1996). Our empirical results give strong and consistent support to the approach that we propose.

## References

Abbott, Thomas (1991)." Producer Price Dispersion, Real Output, and the Analysis of Production." *Journal of Productivity Analysis,* 179-195.

Ackerberg, Daniel, Lanier Benkard, Steven Berry and Ariel Pakes (2007)." Econometric Tools for Analyzing Market Outcomes." *Handbook of Econometrics,* 4171-4276.

Ackerberg, Daniel, Kevin Caves and Garth Frazer (2006)." Structural Identification of Production Functions." *Memo.*

Allen, Franklin (1984). "Reputation and Product Quality." *The RAND Journal of Economics,* 311-327.

Andrews, Donald WK and Biao Lu (2001)." Consistent Model and Moment Selection Procedures for GMM Estimation with Application to Dynamic Panel Data Models." *Journal of Econometrics,* 123-164.

Beaulieu, Joe and Joe Mattey (1999): "The Effects of General Inflation and Idiosynchratic Cost Shocks On Within-Commodity Price Dispersion: Evidence from Microdata." *Review of Economics and Statistics,* 205–216.

Berndt, Ernst and Laurits Christensen (1973)." The Translog Function and the Substitution of Equipment, Structures, and Labor in Us Manufacturing 1929-68." *Journal of Econometrics,* 81-113.

Bils, Mark and Peter Klenow (2004): "Some Evidence on the Importance of Sticky Prices." *Journal of Political Economy,* 947–985.

Clark, Kim and Zvi Griliches (1984). "Productivity Growth And R&D At The Business Level: Results From The Pims Data Base." *R&D, Patents and Productivity, a National Bureau of Economics Research Conference Report,* edited by Zvi Griliches. The University of Chicago Press.

De Loecker, Jan (2011). “Recovering Markups from Production Data.” *International Journal of Industrial Organization,* 350-355.

De Loecker, Jan and Frederic Warzynski, (2012). “Markups and Firm-Level Export Status.” *American Economic Review,* 2437-2471.

Dunne, Timothy and Mark Roberts (1992). “Costs, Demand, And Imperfect Competition as Determinants of Plant-Level Output Prices.” in *Empirical Studies in Industrial Organization: Essays in Honor of Leonard W. Weiss,* edited by David Audretsch and J. Sigfried. Kluwer Academic Publishers.

Foster, Lucia, John Haltiwanger and Chad Syverson (2008). “Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?” *American Economic Review,* 394–425.

Grieco, Paul, Shengyu Li, and Hongsong Zhang (2016). “Production Function Estimation with Unobserved Input Price Dispersion.” *International Economic Review,* 665-690.

Griliches, Zvi and Jacques Mairesse (1998). “Production Functions: The Search for Identification.” In *Econometrics and Economic Theory in the Twentieth Century: The Ragnar Frisch Centennial Symposium.* New York: Cambridge University Press, 169–203.

Griliches, Zvi and Jerry Hausman (1986). “Errors in Variables in Panel Data.” *Journal of econometrics,* 93-118.

Hall, Robert (1986). “Market Structure and Macroeconomic Fluctuations.” *Brookings Papers on Economic Activity*, 285–322.

Hall, Robert (1988). “The Relation between Price and Marginal Cost in U.S. Industry.” *Journal of Political Economy,* 921–47.

Hall, Robert (1990). “Invariance Properties of Solow’s Productivity Residual.” In *Growth/Productivity/Unemployment: Essays to Celebrate Bob Solow’s Birthday,* edited by Peter Diamond. MIT Press.

Hallak, Juan Carlos and Peter K. Schott (2011). “Estimating Cross-Country Differences in Product Quality.” *The Quarterly Journal of Economics,* 417-474.

Hoch, Irving (1955). “Estimation of Production Function Parameters and Testing For Efficiency.” *Econometrica*, 325-326.

Hsieh, Chang-Tai and Klenow, Pete (2009). “Misallocation and Manufacturing TFP in China and India.” *The Quarterly Journal of Economics*, 1403-1448.

Kim, H. Youn (1992). “The Translog Production Function And Variable Returns To Scale.” *The Review of Economics and Statistics,* 546-552.

Kirmani, Amna and Akshay Rao (2000). “No Pain, No Gain: A Critical Review of the Literature on Signaling Unobservable Product Quality.” *Journal of Marketing,* 66-79.

Klette, Tor Jakob and Zvi Griliches (1996). “The Inconsistency of Common Scale Estimators When Output Prices Are Unobserved and Endogenous.” *Journal of Applied Econometrics,* 343-361.

Kokkelenberg, Edward and Sang Nguyen (1989). “Modeling Technical Progress and Total Factor Productivity: A Plant Level Example.” *Journal of Productivity Analysis,* 21-42.

Kugler, Maurice and Eric Verhoogen (2012). “Prices, Plant Size, and Product Quality.” *Review of Economic Studies,* 307–339.

Lichtenberg, Frank and Donald Siegel (1991). “The Impact of R&D Investment on Productivity–New Evidence Using Linked R&D–LRD Data.” *Economic Inquiry,* 203-229.

Mairesse, Jacques and Jordi Jaumandreu (2005). “Panel-Data Estimates Of The Production Function And The Revenue Function: What Difference Does It Make?” *The Scandinavian Journal of Economics,* 651-672.

Marschak, Jacob and William Andrews (1944). “Random Simultaneous Equations and the Theory of Production.” *Econometrica,* 143-205.

Milgrom, Paul and John Roberts (1986). “Price and Advertising Signals of Product Quality.” *Journal of political economy,* 796-821.

Olley, Steven and Ariel Pakes (1996). “The Dynamics of Productivity in the Telecomunications Equipment Industry.) Econometrica, 1263-1297.

Ornaghi, Carmine (2006): “Assessing the Effects of Measurement Errors on the Estimation of Production Functions.” *Journal of Applied Econometrics,* 21(6), 879–891

Rao, Akshay, Lu Qu, and Robert Ruekert (1999). “Signaling Unobservable Product Quality through a Brand Ally.” *Journal of Marketing Research,* 258-268.

Roberts, Mark and Dylan Supina (1996). “Output Price, Markups, and Producer Size.” *European Economic Review,* 909–921.

Roberts, Mark and Dylan Supina (2000). “Ouput Price and Markup Dispersion in Micro Data: The Roles of Producer Heterogeneity and Noise.” *Industrial Organization,* 1-36.

Shapiro, Carl (1982). “Consumer Information, Product Quality, and Seller Reputation.” *The Bell Journal of Economics,* 20-35.

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti (2011). “Growing Like China.” *American Economic Review*, 196-233.

Wooldridge, Jeffrey (2009). “On Estimating Firm-Level Production Functions Using Proxy Variables to Control for Unobservables.” *Economics Letters,* 112-114.

# Appendix A
# Dispersion in Prices

Figure 1 is the histogram of the relative prices of firms from a host of industries. The relative price on the horizontal axis is a firm's price normalized by the yearly industry median price. It is clear from the figure that many firms charge a price substantially different from the yearly industry median price.

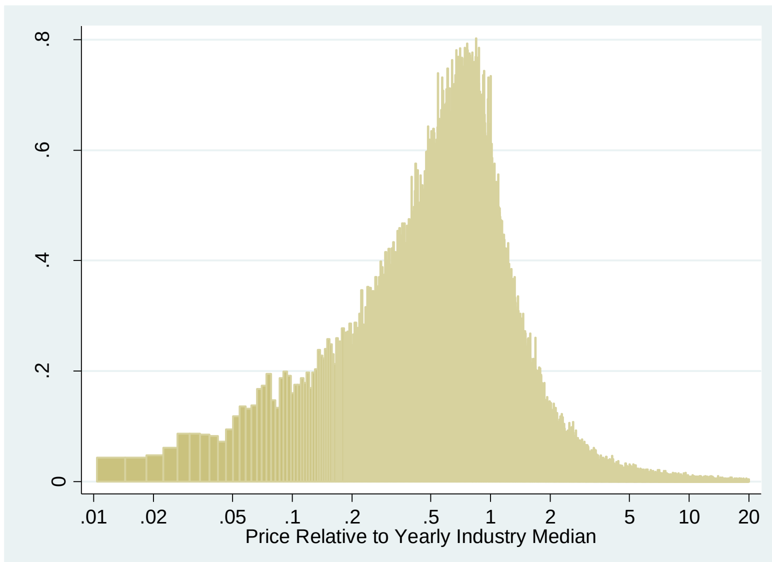


**Figure 1: Histogram of Firm-level Prices Relative to Industry Median**

We find in our data show that firms with the lowest prices ($1^{st}$ percentile) charge a price that is about 9.6 percent of the industry median price, that firms with the highest prices ($99^{t}$ percentile) charge a price that is about 57.4 times of the industry median price. The standard deviation of the logarithm of the relative price is 1.14, which is a very large figure in the logarithm scale for variation. To further make sense of the variation in prices, we compute the Gini Coefficient of the relative prices. Even after we remove observations whose relative prices are below 0.05 or above 20, the Gini Coefficient of prices still stands at 0.526, a number that is close to the income equality in Guatemala in 2014.[11]

Table 19 reports statistics that can reveal the degree of price variations in individual industry after extreme prices are deleted. Columns (2) and (3) show that except for Cement, the cheapest

[11] Guatemala ranks $11^{th}$ among 149 countries in income inequality. Source: CIA Factbook at https://www.cia.gov/library/publications/the-world-factbook/rankorder/2172rank.html.

producers quote prices that are only a small fraction of the industry median price, while the most expensive producers enjoy prices that are several times above the industry median price. Column (4) reports the standard error of the logarithm of prices, which can be interpreted as the variations in the order (i.e., number of digits) of the prices. These figures implies large variations in prices for most of the industries at hand. Column (5) reports the coefficient of variation, showing that the standard error of prices in most industries are greater than the average price.

**Table 19: Price Variations**

| Product | (1) Obs | (2) $5^{th}$ Percentile of Relative Price | (3) $95^{th}$ Percentile of Relative Price | (4) SE of Log Price | (5) SE/Mean |
|---|---|---|---|---|---|
| Bearings | 3029 | .17 | 7.96 | 1.23 | 1.62 |
| Cement | 27016 | .62 | 2.33 | .51 | .91 |
| Ceramics | 3308 | .20 | 9.12 | 1.18 | 1.43 |
| Engineered Wood | 9344 | .41 | 5.10 | .79 | 1.31 |
| Ferroalloys | 4241 | .54 | 3.73 | .65 | 1.35 |
| Furniture | 7708 | .13 | 9.33 | 1.26 | 1.49 |
| Garments | 49605 | .23 | 6.06 | .98 | 1.36 |
| Leather Shoes | 8182 | .25 | 4.11 | .90 | 1.22 |
| Paperware | 10737 | .33 | 8.67 | 1.01 | 1.44 |
| Plastics | 30194 | .39 | 7.46 | .90 | 1.39 |
| Refractories | 4580 | .23 | 6.91 | 1.01 | 1.33 |
| Traditional Medicines | 5178 | .10 | 1.63 | 1.41 | 1.50 |
| Valves | 3309 | .18 | 6.34 | 1.04 | 1.49 |
| All Industries | 166431 | .29 | 6.75 | .99 | 1.37 |

* The relative price is defined as the ratio of a firm's price relative to the industry median price.

* All prices are inflation-adjusted.

* In the last row, the number of observations is the sum of observations in each. All other statistics are the simple average of industry-level values.

To summarize, Table 19 shows that except for highly homogenous product like Cement, there are substantial price variations for most products in our sample. These results confirm the findings on price dispersions reported by Abbott (1991) and others.

# Appendix B
# Production Function and Markup

For the purpose of this section, it suffices to assume a very general function for the physical output:

$$Y_{it} = Y_{it}(X_{it}, \omega_{it}, \epsilon_{it}) \tag{36}$$

where $X_{it} = \left(X_{it}^1, X_{it}^2, \dots X_{it}^V\right)$ are the quantities of the $V$ inputs. Without losing generality, we assume that the first $V-1$ inputs are perfectly variable and the last input is a fixed input.

We assume that the firm minimizes the cost of producing $Y_{it}$ units by choosing variable inputs $\left(X_{it}^1, X_{it}^2, \dots X_{it}^{V-1}\right)$. The auxiliary Lagrangian function of the cost-minimization problem is:

$$\mathcal{L}_{it}\left(X_{it}^1, X_{it}^2, \dots X_{it}^V, \omega_{it}, \lambda_{it}\right) = \sum_{v=1}^{V} P_{it}^v X_{it}^v + \lambda_{it}\left(Y_{it} - Y_{it}(\cdot)\right), \tag{37}$$

where $P_{it}^v$ is the price for input $v$. The Lagrangian multiplier $\lambda_{it} = \frac{\partial \mathcal{L}_{it}}{\partial Y_{it}}$ is the marginal cost of output, $MC_{it}$.

The FOC concerning variable input $v$ is

$$\lambda_{it} = \frac{P_{it}^v}{\partial Y_{it}(\cdot)/\partial X_{it}^v}, v \in \{1,2, \dots, V-1\}. \tag{38}$$

The markup of the firm, $\mu_{it}$, equals

$$\mu_{it} = \frac{P_{it}}{\lambda_{it}} = \frac{P_{it}}{\frac{P_{it}^v}{\partial Y_{it}(\cdot)/\partial X_{it}^v}} = \frac{\partial Y_{it}(\cdot)}{\partial X_{it}^v} \cdot \frac{X_{it}^v}{Y_{it}(\cdot)} \frac{P_{it} Y_{it}(\cdot)}{P_{it}^v X_{it}^v} = \frac{\theta_{it}^v}{P_{it}^v X_{it}^v / P_{it} Y_{it}(\cdot)} = \frac{\theta_{it}^v}{a_{it}^v}, \tag{39}$$

where $\theta_{it}^v = \frac{\partial Y_{it}(\cdot)}{\partial X_{it}^v} \cdot \frac{X_{it}^v}{Y_{it}(\cdot)}$ is the physical output elasticity of input $v$, and $a_{it}^v$ is the expenditure on input $v$ as a share of the firm's revenue.

Equation (39) is derived with the very weak assumption that firms try to minimize their production cost.

# Appendix C
# Output Elasticity of Intermediate Good

With Equation (2), we can show that

$$\theta_{it}^{m}(x_{it},\beta) = \frac{\partial \ln Y_{it}}{\partial m_{it}} = \frac{\partial f(x_{it},\beta)}{\partial m_{it}} - \frac{\partial \ln Q_{it}}{\partial m_{it}}. \tag{40}$$

Given the translog function defined by Equation (9),

$$\frac{\partial f(x_{it},\beta)}{\partial m_{it}} = \beta_m + 2\beta_{mm} m_{it} + \beta_{km} k_{it} + \beta_{lm} l_{it}. \tag{41}$$

The term $\frac{\partial \ln Q_{it}}{\partial m_{it}}$ in Equation (40) is the quality elasticity of intermediate good. If we assume that the product quality of a firm in a period is constant, then we have

$$\theta_{it}^{m}(x_{it},\beta) = \frac{\partial f(x_{it},\beta)}{\partial m_{it}} = \beta_m + 2\beta_{mm} m_{it} + \beta_{km} k_{it} + \beta_{lm} l_{it}. \tag{42}$$

If we would like to follow our assumption on the relation between $Q_{it}$ and $MC_{it}$ closely and define $MC_{it} = \frac{\partial TC_{it}/\partial M_{it}}{\partial Y_{it}/\partial M_{it}}$, we can derive $\theta_{it}^{m}(x_{it},\beta)$ as follows.

Since we assume $Q_{it} = MC_{it} \cdot \exp\big(\zeta(\omega_{it} + \epsilon_{it})\big)$, the quality elasticity of intermediate good is

$$\frac{\partial \ln Q_{it}}{\partial \ln M_{it}} = \frac{\partial \ln\Big(MC_{it} \cdot \exp\big(\zeta(\omega_{it} + \epsilon_{it})\big)\Big)}{\partial \ln M_{it}} = \frac{\partial \ln MC_{it}}{\partial \ln M_{it}}. \tag{43}$$

Note that

$$MC_{it} = \frac{\partial TC_{it}/\partial M_{it}}{\partial Y_{it}/\partial M_{it}} = \frac{P_{it}^{M}}{\partial Y_{it}/\partial M_{it}}. \tag{44}$$

Then we have

$$\frac{\partial \ln Q_{it}}{\partial \ln M_{it}} = \frac{\partial \ln MC_{it}}{\partial \ln M_{it}} = \frac{\partial \ln \dfrac{P_{it}^{M}}{\partial Y_{it}/\partial M_{it}}}{\partial \ln M_{it}} = -\frac{\partial \ln \partial Y_{it}/\partial M_{it}}{\partial \ln M_{it}}. \tag{45}$$

Since $Y_{it} = \exp f(x_{it},\beta)\exp(\omega_{it} + \epsilon_{it})$,

$$
\begin{aligned}
\partial Y_{it}/\partial M_{it} &= \frac{\partial \exp f(x_{it},\beta)\exp(\omega_{it}+\epsilon_{it})}{\partial M_{it}} \\
&= \frac{\partial f(x_{it},\beta)}{\partial M_{it}}\exp(\exp f(x_{it},\beta))\exp(\omega_{it}+\epsilon_{it}) \\
&= \frac{\partial f(x_{it},\beta)}{\partial M_{it}} Y_{it} \\
&= \frac{\partial f(x_{it},\beta)}{\partial m_{it}}\frac{\partial m_{it}}{\partial M_{it}} Y_{it} \\
&= f_m(x_{it},\beta)\frac{Y_{it}}{M_{it}}.
\end{aligned} \tag{46}
$$

Then

$$
\begin{aligned}
\frac{\partial \ln Q_{it}}{\partial \ln M_{it}} &= -\frac{\partial \ln \partial Y_{it}/\partial M_{it}}{\partial \ln M_{it}} = -\frac{\partial \ln\left(f_m(x_{it},\beta)\frac{Y_{it}}{M_{it}}\right)}{\partial \ln M_{it}} \\
&= -\frac{\partial \ln\left(f_m(x_{it},\beta)\right)}{\partial \ln M_{it}} - \frac{\partial \ln Y_{it}}{\partial \ln M_{it}} + \frac{\partial \ln M_{it}}{\partial \ln M_{it}} \\
&= -\frac{\partial \ln(\beta_m + 2\beta_{mm}m_{it} + \beta_{km}k_{it} + \beta_{lm}l_{it})}{\partial m_{it}} - \frac{\partial \ln Y_{it}}{\partial \ln M_{it}} + 1 \\
&= -\frac{2\beta_{mm}}{f_m(x_{it},\beta)} - f_m(x_{it},\beta) + 1.
\end{aligned} \tag{47}
$$

Plug equations (47) into Equation (40):

$$
\begin{aligned}
\theta_{it}^{m}(x_{it},\beta) &= \frac{\partial \ln f(X_{it},\beta)}{\partial m_{it}} - \frac{\partial \ln Q_{it}}{\partial M_{it}} \\
&= f_m(x_{it},\beta) - \left(-\frac{2\beta_{mm}}{f_m(x_{it},\beta)} - f_m(x_{it},\beta) + 1\right) \\
&= 2f_m(x_{it},\beta) + \frac{2\beta_{mm}}{f_m(x_{it},\beta)} - 1
\end{aligned} \tag{48}
$$

$$
= 2(\beta_m + 2\beta_{mm}m_{it} + \beta_{km}k_{it} + \beta_{lm}l_{it}) + \frac{2\beta_{mm}}{\beta_m + 2\beta_{mm}m_{it} + \beta_{km}k_{it} + \beta_{lm}l_{it}} - 1
$$