On the estimation of marginal cost

Delis, Manthos D and Iosifidi, Maria and Tsionas, Efthymios

1 December 2012

On the estimation of marginal cost

Manthos D. Delis*

The Surrey Business School, Faculty of Business, Economics and Law, University of Surrey, Guildford, GU2 7XH, UK

Maria Iosifidi

Department of Economics, City University, Northampton Square, London EC1V 0HB, UK
and
Athens University of Economics and Business, 76 Patission St., Athens, 10434, Greece

Efthymios G. Tsionas

Department of Economics, Athens University of Economics and Business
76 Patission Street, 10434 Athens, Greece

This version: July 20, 2012

We are grateful to Severin Borenstein, James Bushnell, and Frank Wolak for providing their data from the California's electricity market and valuable comments. We are also grateful to Iftekhar Hasan, Sofronis Clerides and Eugenia Vella for many helpful comments and suggestions, as well as to seminar participants at Cass Business School, Brunel University, Surrey Business School and Athens University of Economics and Business. Delis acknowledges financial support from City University and University of Surrey and Tsionas from the Athens University of Economics and Business. The usual disclaimer applies.

* Corresponding author. E-mail address: elmanthos@hotmail.com (M.D. Delis)

On the estimation of marginal cost

Abstract

This article proposes a general empirical method for the estimation of marginal cost of individual firms. The new method employs the smooth coefficient model, which has a number of appealing features when applied to cost functions. The empirical analysis uses data from a unique sample from which we observe marginal cost. We compare the estimates from the proposed method with the true values of marginal cost, and the estimates of marginal cost that we obtain through conventional parametric methods. We show that the proposed method produces estimated values of marginal cost that very closely approximate the true values of marginal cost. In contrast, the results from conventional parametric methods are significantly biased and provide invalid inference.

Keywords: Estimation of marginal cost; Parametric models; Smooth coefficient model; Actual and simulated data

JEL classification: C14; C81; D24; G21; Q40

1. Introduction

Since the contribution of Alfred Marshall, one of the most fundamental and widely used concepts in economics and management sciences is marginal cost. However, in the great majority of industries, the marginal cost of firms cannot be readily observed in the data. Thus, researchers and practitioners have to rely on estimates that might or might not be robust. This problem creates notorious difficulties in analyzing empirically basic microeconomic theories on industrial organization and management. These difficulties also extend to policy-makers with concern to the cost structures of firms and industry conduct. In this paper, we revisit the issue of the estimation of marginal cost with a standard cost function and show how to derive robust estimates for individual firms in our sample and for each point in time (i.e., at the observation level). The new method is quite general, does not rely on strong assumptions, and can be applied to any industry.

In particular, we propose a method that relies on the estimation of a standard cost function that uses a nonparametric estimation model, namely the smooth coefficient model. Nonparametric methods are quite flexible and allow obtaining observation-specific estimates of the parameters of a basic cost function and, in turn, of marginal cost. In other words, we can obtain marginal cost estimates at the firm-level and for each point in time by using the standard textbook principle of the derivative of total production cost over the firm's output. Further, this approach allows the relaxation of a number of restrictive assumptions that pertain to the estimation of cost functions; the most important one being the assumption of a specific functional form that is required under parametric methods. Reiss and Wolak (2007), among many others, are skeptical about this assumption, because the structure of the cost and output data can bias marginal cost estimates to an unknown magnitude and direction. In contrast, within the smooth coefficient model, variables can vary according to information derived from the data, and this represents a closer approximation of reality. Also, this flexibility allows the researcher to use large international samples of firms without being concerned that certain industries in different countries or firms within one industry face or adopt different technologies.

Even though the smooth coefficient model is theoretically well established in the statistics literature, researchers have never verified that the parameter estimates approximate the true values of the parameters. Because of the lack of studies on this front, researchers are skeptical about the applicability of this method to the estimation of marginal cost. To confirm that the estimated values of marginal cost from the smooth coefficient model approximate the true marginal cost well, we conduct a number of empirical tests. Specifically, we use data and implications from important studies on one of the few industries where the true values of marginal cost can be readily observed from the available data and estimated from the same data set: California's electricity industry. We then compare the values of the true marginal cost to the values of the estimated costs and show that the correlation between the two is very high and that their distribution densities are quite similar. We carry out these tests using two data sets: actual aggregate data for all firms in the industry and simulated panel data that comprise information for each firm at each point in time.

In addition, we estimate the same total-cost equation using state-of-the-art parametric models. In particular, we use the two data sets and impose either a log-linear or a flexible translog parametric form to the cost function. We find that estimates of the marginal cost using the parametric methods are worse approximations of the values of the true marginal cost compared to the nonparametric method. For the panel data set where different firms have different production technologies, the bias is so large that estimates of marginal cost might lead to completely invalid inferences in a wide array of applications related to industrial-organization theory, competition policy, applied microeconomic projects, and management science studies.

Empirical studies of marginal cost go back at least to Rosse (1967). In virtually all of these studies, researchers carry out the estimation of marginal cost by using parametric econometric methods as well as the assumption that the production technology is the same among firms in the same industry or between different groups of firms. Therefore, based on our findings from the empirical tests, the results of these studies might be considerably biased. The realization of this problem, among other issues, leads some researchers to avoid estimating cost equations and to rely on alternative empirical techniques to infer firm behavior and industry conduct (e.g., by using demand equations, as in Nevo, 2001). Other important contributions utilize data from the few industries where researchers can observe the true marginal cost. For example, Wolfram (1999), Borenstein et al. (2002), and Fabrizio et al. (2007) use such data from the electricity industry and Genesove and Mullin (1998) from the sugar industry.

A few relatively recent studies use smooth coefficient models to estimate cost or production functions, but these studies have different objectives from ours and between themselves. Kumbhakar et al. (2007) use a local maximum-likelihood technique to estimate a stochastic production frontier with the aim of obtaining efficiency estimates for various firms. Kumbhakar and Tsionas (2008) use a similar method to estimate a stochastic cost frontier with a related objective. Hartarska et al. (2010) and Asaftei et al. (2008) use a smooth coefficient model to estimate economies of scope in microfinance lending and in banking respectively. To the best of our knowledge, our study is the first that uses the smooth coefficient model to obtain marginal cost estimates and prove that these estimates are very close approximations of the true values of marginal cost.

The rest of the paper proceeds as follows. In Section 2, we discuss the data set of Borenstein et al. (2002) from the California's electricity market and the panel data set that we generate by using the implications of Borenstein et al. (2002) and Kim and Knittel (2006). The two data sets contain both true (observed) values of marginal cost and the data required to estimate marginal cost from the cost function. In Section 3, we use the two data sets to show that conventional parametric methods fail to approximate marginal cost to a high degree. In Section 4, we reestimate marginal cost by using the smooth coefficient model and a novel strategy and show that this new method produces much more accurate estimates of the true marginal cost for both the actual and the simulated sample. Section 5 concludes the paper.

2. Data

In this section, we provide information on the setup of the empirical tests to identify whether the values of the estimated costs approximate the true marginal cost. In particular, we describe two data sets. The first is an actual data set that comprises information on aggregate data across firms from California's electricity industry. The second is a simulated panel data set that uses information from the actual data set and a general cost function.

2.1. The actual data

The most important requirement for the estimation of the marginal cost of each firm at each point in time is to guarantee that the estimated values approximate the true values of marginal cost fairly well. The only way to examine whether this approximation holds is to carry out an empirical test that uses information from an industry in which the data required to estimate the cost function and derive the marginal cost are readily available. Then, we can compare the true values of marginal cost with the estimated costs by using basic statistical methods and infer whether the estimates approximate the true marginal cost fairly well.

However, even finding an industry in which the true marginal cost is observed and that also has the available data to estimate the marginal cost is not an easy task. Among the very few candidates, the electricity industry is probably the best. A few well-known studies (e.g., Wolfram, 1999; Borenstein et al., 2002; Kim and Knittel, 2006) with somewhat different objectives among them, use the electricity industry precisely because marginal cost can be calculated directly from the data.

Using the same data set as Borenstein et al. (2002) from the California electricity market might be ideal. These authors use hourly data from electricity generation firms operating in the California electricity market over the period of 1998 to 2000. Unfortunately, the data from Borenstein et al. (2002) only provide the required information at the aggregate level across all firms in the industry for each hour. Thus, their data are time-series data. Because these data provide information on the true marginal cost and on the variables needed to estimate it from the cost function, they are still suitable to examine the hypothesis that the true marginal cost ingthe estimated one. The data set comprises 21,217 observations on marginal cost (*mc*), total cost (*tc*), industry output (*Q*), and the input prices of production. Most of the firms in the industry essentially use two input prices, namely the average daily price of natural gas for California (*png*) and the price of NOx permits (*pno*). These are the input prices that Kim and Knittel (2006) also use. We report summary statistics for these variables in Panel A of Table 1.

[INSERT TABLE 1]

2.2. The simulated panel data

Notably, most of the research studies that estimate marginal cost from the cost function use panel data from firms operating in a particular industry. The difference between studies that use panel data and studies that use aggregate data can be considerable if different firms have different production technologies and cost structures. This difference implies that imposing parametric structures on the cost function can yield larger biases in panel-data studies. To examine the hypothesis that the true marginal cost equals the estimated one using panel data, we generate simulated-panel data from the underlying information that the actual data provides.

In particular, we carry out a data generation process for the variables *q* (output of firm *i*), *png*, *pno,* and *mc*. The dimension of the panel is set to 4,000 hourly observations times five firms that equal 20,000 observations in total.[1] To generate data for these variables, we use a chi-squared distribution and the mean values from the actual data set. We favor the chi-squared distribution over, for example, the log-normal because of the shape of the distribution of the actual data and the fact that cost data are usually positively skewed. We define *q* as exogenous to the cost equation, which has important implications for the validity of the estimation procedure below. We also generate data on *mc* in a way that *tc* has a bimodal distribution. We favor this method to pose a more stringent requirement on the parametric and nonparametric models to be studied later. The requirement is that the estimates of marginal cost must be able to approximate the true marginal cost even though total cost is clustered in two modes. This clustering could arise, for example, from an industry where there are two groups of firms such as a number of large firms with high costs and a number of small ones with lower costs or from an industry where there is seasonality in the cost data with a period of high costs and a period of low costs.

For the data generation process of *mc*, we also use an alternative to specify the functional form. This alternative needs to be the most general data generation process available; because we do not want to give any parametric functional form of the cost equation, such as the log-linear or the

[1] We also experiment by generating 21,217 observations times five firms, i.e., 106,085 observations in total. However, the larger number of observations in this panel increases the computational burden of the semi-parametric analysis in the software Gauss, which is already high even for a powerful 64-bit pc, without any effect on the final results. The main results for the 21,217x5 panel are available on request.

translog, an unfavorable advantage over the other. Therefore, we also experiment with *mc* generated from regression estimates that use the actual data and (i) a log-linear specification for the total cost equation or (ii) a translog specification of the total cost equation. Not surprisingly, in the first case the best fit to the simulated *mc* are the estimates of *mc* from the log-linear specification, while in the second case the best fit are the estimates of *mc* from the translog specification. Consequently, these tests are not very informative if one wants to examine cases where the underlying structure of the data is either more complex or unknown.

We provide summary statistics for the generated variables in Panel B of Table 1. From this data set, we calculate total costs for each available observation from the textbook formula:

$$mc = \Delta tc / \Delta q \Rightarrow mc = (tc_1 - tc_0)/(q_1 - q_0) \Rightarrow tc_1 = mc(q_1 - q_0) + tc_0 . \qquad (1)$$

This is a good approximation for *tc*, because of the high frequency of the data.[2]

3. Estimation of marginal cost with existing parametric methods

Before discussing the nonparametric method to estimate the total cost equation, we use an existing parametric method and the implications of the extensive literature on this issue to obtain estimates of marginal cost.[3] We show that the parametric methods fail to approximate marginal cost to a reasonable degree. This failure raises serious doubts about the findings and implications of the literature on several microeconomic, industrial-organization and management-science projects.

We closely follow the conventional approaches to the estimation of marginal cost (e.g., Koetter et al., 2011) and specify either a log-linear or a flexible translog cost function. The log-linear cost function (log of Cobb-Douglas) takes the form of

$$\ln(tc / pno) = c_0 + c_1 \ln(png / pno) + c_2 \ln q + u , \qquad (2)$$

and the translog the form of

---

[2] In most industries total cost data are generally available. The problem is that if the frequency of the data is, e.g., annual, it cannot be use the textbook formula to infer the actual marginal cost because of changing economic conditions (e.g., technology, macroeconomic environment, etc.).

[3] There are literally hundreds of studies with various objectives that estimate the marginal cost of firms. Only indicatively, see Hall (1988), Bresnahan (1989), Roeger (1995), Berg and Kim (1998), Konings et al. (2005) and Koetter et al. (2011).

$$\ln(tc / pno) = c_0 + c_1 \ln(png / pno) + c_2 \ln q + c_3 \ln(png / pno) * \ln q + c_4 1/2 * (\ln(png / pno))^2 + c_5 1/2 * (\ln q)^2 + u. \quad (3)$$

In these specifications, we drop for simplicity the subscripts *i* and *t* from the variables. Further, we impose homogeneity of degree one on input prices by dividing *png* and *tc* by *pno*. From Equations (2) and (3), respectively, marginal cost can be calculated for each observation *it* as

$$mc = tc / q * c_5 \quad (4)$$

$$mc = tc / q[c_2 + c_3 \ln(png / pno) + c_5 \ln q]. \quad (5)$$

The literature proposes the estimation of the cost equation using either least-squares-based methods or maximum likelihood (ML). We first estimate Equations (2) and (3) by using the actual data and OLS. As the values of the variables added to these equations vary, this method in fact yields observation-specific estimates of *mc*. Of course, this is a crucial issue underlying the shape of the values of the estimated *mc*. We report coefficient estimates and *t*-statistics in the first two columns of Table 2. All estimated coefficients are statistically significant at the 1% level and bear the expected sign. Also, the high values on the adjusted R-squared show that very little is left unexplained.

[INSERT TABLE 2]

In these specifications we assume that output is exogenous to the cost equation. Reiss and Wolak (2007) have a very nice discussion on these issues and note correctly that the exogeneity of output and input prices in cost functions depends on the type of industry. Specifically, endogeneity of *q* arises if, for example, the cost function comes from a production function $q_i = f(A_i, L_i, K_i)$ that includes a component A_i characterizing technology as different among firms. Then, the cost equation will also be a function of *A* and if *A* is unobserved it renders *q* endogenous.[4] If this is the case, there are two choices. The first is to identify reasonable estimates of A_i and add them to the estimated equation, and the second is to instrument *q*. Of course, the results of the empirical tests (i.e., the extent

[4] Reiss and Wolak (2007, pp. 28-30) give a clear example of how this endogeneity emerges for the Cobb-Douglas technology. The exogeneity of input prices is not an issue within the smooth coefficient model; because the nonparametric regression compares only across observations who have the same values for the rest of the explanatory variables (here all variables besides *q*) and differ only in the treatment variable (here *q*), whereas the parametric regression combines all observations in a single "global" regression (Frolich, 2008; Blundell and Dias, 2008). Thus, endogeneity of prices does not affect the estimated coefficients on *q* in the nonparametric regressions. This is an additional advantage to the smooth coefficient model.

to which the estimated marginal cost approximates the true marginal cost) can solve this potential identification issue. Further, the technology of different firms in the electricity industry should not be widely different, and instrumental variables regressions are less efficient than OLS if the endogeneity bias is small. So, in the last two columns of Table 2 we report the results from the estimations of Equations (2) and (3) using a two-stage least squares (2SLS) estimation method and the changes in the two input prices as instruments. In this setting, prices are exogenous (see Borenstein et al., 2002) and, therefore, their changes qualify as proper instrumental variables.

Using Equations (4) and (5), we obtain the marginal cost estimates for the four different specifications in Table 2. We report summary statistics for these estimates in Table 3 that show the mean values from the different specifications are very close to the mean value of the true marginal cost. In this respect, the parametric models approximate the true marginal cost fairly well. The problems start with the examination of the correlation coefficients between the parametric estimates of marginal cost and the true values of marginal cost (see Table 4). The higher correlation is between the marginal cost obtained from the estimation of the translog with OLS (mc_{trols}) and the true marginal cost (*mc*) and is equal to 0.902. The rest of the relevant correlation coefficients are below 0.9.

[INSERT TABLE 3]

[INSERT TABLE 4]

Far more importantly, the comparison of the probability density function (pdf) of the estimated values of marginal cost with the equivalent of the values of the true marginal cost (shown on Figures 1–4) divulges important differences. The first two figures show that the estimation of the cost function with OLS somewhat overestimates the total cost for a significant number of observations. Further, the results from both the log-linear and the translog specifications illustrate that 2SLS produces marginal cost estimates that deviate from true marginal cost to a larger degree. Specifically, the correlation coefficients between mc and mc_{ll2sls} or mc_{tr2sls} are lower compared to the equivalent ones from the OLS. Most importantly, the pdf of mc_{ll2sls} and mc_{tr2sls} show a worse fit to the true marginal cost compared to the OLS equivalents (see Figures 3 and 4). All these are clear evidence that any use of these estimates for research or policy purposes biases inference, even for an industry with very simple technology such as the one described by this data set.

[INSERT FIGURES 1-4]

Subsequently, we use the simulated data that are generated in a way that poses more stringent requirements on the parametric models. In Table 5, we report the results from applying OLS to the log-linear and the translog specifications. Both the coefficients on the price and the quantity variables added to the log-linear specification and the "actual" effect of these variables in the translog specification (when taking partial effects) are positive and statistically significant at the 1% level. We proceed by estimating *mc* from each specification, given Equations (4) and (5). The mean values of the estimated marginal cost, reported in Table 6, are quite higher than the mean value of the actual marginal cost (reported in Panel B of Table 1). This shows that, on average, both parametric specifications overestimate marginal cost in the sample. More importantly, the correlation coefficient between the estimates of marginal cost obtained from the log-lineal specification (mc_{llols}) and the actual values of marginal cost is nearly zero (see Table 7). The equivalent correlation coefficient involving the estimates from the translog is better, but still as low as 0.342.

[INSERT TABLE 5]

[INSERT TABLE 6]

[INSERT TABLE 7]

The reason behind these findings is the fact that the pdf of *tc* is bimodal because of the way we generated the data. In Figure 5, we present the pdf of mc_{llols} versus the pdf of the actual *mc*. Even though the pdf of *tc* is bimodal, the pdf of the true $mc = \Delta tc/\Delta q$ is unimodal. In contrast, the pdf of mc_{lolls} is bimodal that is imposed by the appearance of *tc* in Equation (4) and no other terms that improve the flexibility of the estimates. Therefore, the estimates of marginal cost from the log-linear specification can be severely biased if the structure of the cost data is unfavorable. As Figure 6 shows, the flexibility of the estimates is significantly improved when using the translog. Yet, the fit of the estimates to the actual *mc* is still far from optimal, with the estimates of *mc* displaying a much higher standard deviation (see also relevant values in Tables 1 and 6).

[INSERT FIGURE 5]

[INSERT FIGURE 6]

We carry out extensive sensitivity analyses on these findings by (i) estimating Equations (2) and (3) with a ML method, (ii) using other functional forms like the generalized Leontief, and (iii) removing the outliers of marginal cost. The results from these sensitivity analyses do not improve on the results already presented. The main component driving the findings is the presence of the actual variables in Equations (4) and (5), which are imposed by the parametric assumptions. Indeed, one can experiment with other variants of parametric models and perhaps find a specification that yields estimates that better approximate the true values of marginal cost for the present sample. However, in the vast majority of the industries, the researcher or the practitioner does not know the true marginal cost to compare her estimates with it. This is why she proceeds with the estimation. Here we provide a general method that approximates the true marginal cost quite well, without relying on specific assumptions on the functional form, the specification of the marginal cost and the type of the industry.

4. Estimation of marginal cost with a nonparametric method

This section presents a new method for the estimation of marginal cost at each point (observation) in the data. We contend that the method is general enough to be applied to any industry with a sufficient number of observations in which the true marginal cost is not observed. The reason for this is that the method does not rely on any specific parametric assumptions about the type of technology used in the industry or by different firms within the industry and the extent of the market. The only requirements are: (i) the econometric methodology is robust to a reasonable degree and (ii) the variables employed do not have measurement error.

The new method relies on the estimation of a total cost function by using the well established semi-parametric or nonparametric smooth coefficient model. Given the theoretical foundation of this class of models, the choice of the functional form should not play any role in the values of the estimated coefficients and, thus, the most basic specification possible can be used. Therefore, we rely on the estimation of a general linear cost function in the form of

$$tc_{it} = a_0 + a_1 png_{it} + a_2 pno_{it} + a_3 q_{it} + e_{it} \,. \qquad (6)$$

For this cost equation a_4 is equal to the marginal cost. Note that the smooth coefficient model is a varying-coefficient model because the model allows as many estimates for a_4 as there are observations in the data. Thus, we obtain direct estimates of the marginal cost from a_4 and compare them with the true values of marginal cost.

4.1. The smooth coefficient model

We utilize a semi-parametric smooth coefficient model according to Fan (1992) and Mamuneas et al. (2006) that follows the local polynomial regression of Stone (1977). The analysis in our method closely follows Mamuneas et al. (2006) who investigate the impact of human capital on economic growth. For more information on this approach refer to, for example, Hoover et al. 1998, Fan and Zhang (1999), Cai et al. (2000), and Li et al. (2002).

Our initial assumption is that the data are $\{Y_i, W_i\}$, $i = 1,\ldots, n$, where Y is the response variable that equals *tc* and W is the matrix of the independent variables that comprises *png*, *pno*, and q. We suppress the subscript t that reflects hours. Also, we have $W_i = \{X_i, V_i\}$, where X comprises the input prices *png* and *pno*, while V the terms on q. We can now rewrite Equation (6) in econometric form as

$$Y_i = E(Y_i \mid W_i) + e_i = X_i\beta_1 + V_i\beta_2(Z_i) + e_i\,. \tag{7}$$

In Equation (7), β_2 is a function of one or more variables with dimension k added to the vector Z, which is an important element of the analysis and will be discussed below. The presence of a linear part in Equation (7) is in line with the idea of the semi-parametric model as opposed to a fully nonparametric model (e.g., Zhang et al., 2002). The coefficients of this part are estimated in a first step as averages of the polynomial fitting by using an initial bandwidth chosen by cross-validation (Hoover et al., 1998). In the second step we use these average estimates to redefine the dependent variable as

$$Y_i^* = V_i\beta_2(z) + e_i^*\,, \tag{8}$$

where the stars denote the redefined dependent variable and error term.

The coefficient $\beta_2(z)$ that is evaluated at a z point of Z is a smooth but unknown function of z.[5] Here, we estimate $\beta_2(z)$ using a local least squares approach of the form

$$\begin{aligned}\hat{\beta}_2(z) &= \left[(n\lambda^k)^{-1}\sum_{j=1}^{n} V_j^2 K\left(\frac{z_j - z}{\lambda}\right)\right]^{-1}\left[(n\lambda^k)^{-1}\sum_{j=1}^{n} V_j Y_j^* K\left(\frac{z_j - z}{\lambda}\right)\right] \\ &= [B_n(z)]^{-1} C_n(z)\end{aligned} \tag{9}$$

where $B_n(z) = (n\lambda^k)^{-1}\sum_{j=1}^{n} V_j^2 K\left(\frac{z_j - z}{\lambda}\right)$, $C_n(z) = n\lambda^k)^{-1}\sum_{j=1}^{n} V_j Y_j^* K\left(\frac{z_j - z}{\lambda}\right)$.

In Equation (9), $K(z, \lambda)$ is a kernel function and λ is the smoothing parameter for sample size n.

The reasoning behind the local least squares estimator in Equation (9) is as follows. If we assume that z is a scalar and K is a uniform kernel, then Equation (9) can be written as

$$\hat{\beta}_2(z) = \left[\sum_{|z_j - z| \le \lambda} V_j^2\right]^{-1}\left[\sum_{|z_j - z| \le \lambda} V_j Y_j^*\right]. \tag{10}$$

In Equation (10), $\hat{\beta}_2(z)$ is a least squares estimator obtained by regressing Y_j^* on V_j, using the observations of $(V_j,\ Y_j^*)$ for which the corresponding z_j is close to z, that is, $|z_j - z| \le \lambda$. Therefore, to estimate $\hat{\beta}_2(z)$ we only use observations within this "sliding window". Note that no assumptions are made about this estimator globally, but locally, i.e. within the sliding window, we assume that $\hat{\beta}_2(z)$ can be well-approximated. Also, because $\hat{\beta}_2(z)$ is a smooth function of z, $\lambda \to 0$ is small when $|z_j - z|$ is small. The condition that $n\lambda^k \to \infty$ is large ensures that we have sufficient observations within the interval $|z_j - z| \le \lambda$ when $\beta_2(z_j)$ is close to $\beta_2(z)$. Therefore, under the conditions that $\lambda \to 0$ and $n\lambda^k \to \infty$, the local least squares regression of Y_j^* on V_j provides a consistent estimate of $\beta_2(z)$. This is why the estimation method is usually referred to as a local regression.

A critical issue in the estimation process is the choice of the variable(s) to comprise Z. The best candidates are variables that are highly correlated with β_2 but that also allow variation for β_2

[5] Mamuneas et al. (2006) discuss in detail how this function can take specific parametric formulations (such as linear) that can be tested against the general unknown specification.

across firms and time. In a cost function, the natural candidates to use are the input prices as *Z*. The advantage of this choice is that input prices most certainly affect β_2 to a large extent. This has been shown many times when researchers employ a translog specification, which includes multiplicative terms of output with input prices, to estimate the cost function parametrically. In our sample, specification (2) of Table 2 provides evidence that this is indeed the case. Thus, we primarily use the linear combination of *png* and *pno* as *Z*, but we also experiment with a linear combination of both input prices and find no significant changes in the results.

Estimation of Equation (6) using the aforementioned technique presents some important interrelated advantages besides that of obtaining observation-specific estimates of marginal cost through localization. By definition, no assumption regarding the functional form of the underlying production relation is needed, and researchers commonly have difficulty in being certain that they are choosing the "correct" functional form. This difficulty implies that not all firms necessarily have the same production function and, therefore, the observation-specific estimates of marginal cost are free of such misspecification bias. In other words, estimates vary according to the information derived from the data, which represents a closer approximation to reality. This flexibility allows the researcher to use large international samples without being concerned that certain industries in different countries or firms within one industry face different technologies. Further, economic hypotheses are not rejected because researchers choose an "improper" functional form. For these reasons, recent literature has used similar nonparametric techniques in a variety of economic problems, including those involving the estimation of cost functions (see, e.g., Kumbhakar et al. 2007).

Related to this discussion of the smooth coefficient model and the estimation procedure are three critical issues. First, semiparametric and nonparametric techniques have to be applied to large data sets to avoid the so-called "curse of dimensionality." If observations within the interval $|z_j - z| \leq \lambda$ where $\beta_2(z_j)$ is close to $\beta_2(z)$ are insufficient, then results are likely to be biased. Fortunately, this is not an issue for most micro-level studies like the present one, where datasets are quite large. Second, any local regression might be highly sensitive to outliers and, therefore, a researcher should be cautious when appropriately excluding extreme values from the data set. We tackle this issue by means of examining the robustness of our results for a trimmed sample without outliers.

Third, the majority of the literature on the estimation of cost functions uses either least squares or ML methods. Least squares methods are consistent under the assumption that all right-hand side variables are exogenous. The ML method is useful when the researcher needs to estimate some model of productive or technical efficiency (see Kumbhakar and Lovell, 2000), but still relies on the same exogeneity assumption. As discussed in Section 2, the output variable q can be endogenous in the cost equation even though the majority of the literature on the estimation of cost functions disregards this element. Also, in Section 2, we established that in this data set the OLS-based methods are optimal.

4.2. Empirical findings

In this section, we report and analyze the estimation results from Equation (6) of the smooth coefficient model and compare these results with the true marginal cost that the analysis in Section 2 provides. We demonstrate that the two are almost equal, which shows that the new method provides improved estimates of marginal cost compared to the equivalent ones from the parametric models.

We start by applying the smooth coefficient model and local regression to the actual data. In Panel A of Table 8, we report summary statistics for the coefficient estimates of *png*, *pno*, and q. The average of the estimated coefficients for q, that is, the marginal cost (denoted as mc_{scm}), is 35.10 that is very close to the average of the true marginal cost (34.45). Similarly, the correlation coefficient between mc_{scm} and the true *mc*, reported in the last row of Panel A, is as high as 0.974. Further, in Figure 7, we graph the pdf of mc_{scm} against that of the true *mc* as we did for the parametric models earlier. The figure shows a much better fit than before with the estimated marginal cost mapping the true marginal cost almost perfectly. This is first-hand evidence that the proposed method provides superior estimates of the marginal cost at each point in the data compared to the estimates from the parametric models.

[INSERT TABLE 8]

[INSERT FIGURE 7]

To guarantee that the smooth coefficient model can produce improved estimates of marginal cost even if we impose more stringent structure on the cost data, we rerun the method with the simulated panel data set. We report averages of the coefficient estimates in Panel B of Table 8. Again, the mean of the estimates on mc_{scm} is very close to the mean of the true *mc* from the simulated panel (34.7 versus 34.4). Most importantly, the correlation coefficient between the two is as high as 95%, which is a great improvement compared to the equivalent one from the translog equation of 34%. This improvement shows that if the marginal cost estimates are used to, for example, calculate price-cost margins, then the accuracy of the margins will improve considerably if the marginal cost is estimated through the nonparametric method. Further, we compare the pdf of mc_{scm} obtained from the simulated data set to the true *mc* in Figure 8. Again, even though not perfect in this case, the mapping of the values of mc_{scm} against those of the true *mc* improves significantly compared to Figure 6.

[INSERT FIGURE 8]

We conduct a number of sensitivity analyses on the results of this section. In particular, (i) we use only *png* as *Z*, instead of the linear combination of both input prices, (ii) we obtain estimates on *mc* by trimming 1% of outliers from both edges of the distribution of *mc*, and (iii) we include fixed effects and time effects among the regressors. The results, not reported here but available on request, are not significantly different from those reported in Table 8 and shown on Figures 7 and 8.

As a final exercise, we estimate separately the impact of all the different estimates of marginal cost obtained with parametric or nonparametric methods on the true marginal cost to acquire the R-squared statistic of the regression. We report the results for the five regressions in Table 9. Evidently, the coefficients on the marginal cost estimates from the smooth coefficient model (reported in columns 7 and 8) are close to unity, while the values of the R-squared statistics are the highest ones among those that Table 9 reports. This test again indicates the superiority of our nonparametric method to estimate marginal cost.

[INSERT TABLE 9]

Our results show that the proposed method for the estimation of marginal cost represents a great improvement over the current estimation methods. The correlation coefficients in Table 8 are so high that inference based on the new estimates can be carried out with confidence. For example, if one has information on prices of products and needs marginal cost estimates to calculate price-cost margins, then the estimates from the new method will approximate the true price-cost margins in a much better way than the estimates from the conventional methods. Many other examples can be provided from the economics and business literature where estimates of marginal cost are required for research and policy reasons. Further, the very good fit of the estimated values of *mc* to the true values show that in the actual data set the endogeneity of q is not much of an issue and confirms that least-squares based methods are suitable for the analysis of this data set.[6]

5. Conclusions

Research and policy have long been built on marginal cost estimates obtained from empirical models that impose strong parametric assumptions on cost functions. The objective of this paper is to evaluate these methods and, most importantly, to propose a better alternative. We first use actual data for marginal cost and for variables required to estimate the marginal cost with a cost function. These data are from the well established empirical industrial-organization literature that analyzes the electricity market, which is a unique case for observing marginal cost directly. Using information from this data set, we also generate a panel data set from a suitable simulation process. Subsequently, we use both data sets to estimate the cost equation with (i) conventional parametric methods and (ii) with a new method that proposes the estimation of the cost function with nonparametric techniques.

We first show that conventional methods provide inaccurate estimates of marginal cost when applied to both the actual and the simulated data sets. The bias can be quite large and invalidate economic inference to a considerable extent. Next, we estimate the cost equation using the nonparametric smooth coefficient model. This framework has a number of appealing features, the most important one being that the framework allows us to obtain coefficient estimates of marginal cost equal to the number of observations without imposing a specific functional form on the cost equation. We show that the new method produces estimates of marginal cost that approximate the true marginal cost very closely. In particular, (i) the correlation coefficient between the true and the estimated marginal cost is equal to or higher than 95% in all alternative specifications that we estimate, (ii) the

[6] This is also the case for the simulated data set in which we define q as exogenous.

probability density function of the estimated marginal cost maps very closely the one of the true marginal cost and (iii) the R-squared of the regression of true marginal cost on the estimated one gives a very high R-squared statistic. We contend that these findings call for a reconsideration of the literature that relies on marginal cost estimates with a basis of parametric assumptions. This is of special relevance to the literature involving identification of industry conduct.

Further work needs to be done on the estimation of cost equations that use smooth coefficient models for smaller data sets that are prone to the so-called "curse of dimensionality." Recent econometric literature proposes various methods to overcome this problem. Overcoming this problem would also allow re-evaluation of marginal cost measures within New-Keynesian macroeconomic models, which is another source of debate and friction within the economics literature, because of the requirement to estimate marginal cost for the aggregate economy. Moreover, the framework of this paper can be used to reexamine elements of profit-maximization by equating estimates of marginal cost and marginal revenue. Finally, the present analysis needs to be augmented by proposing a consistent estimator of marginal cost when the firm's output is endogenous. This augmentation calls for an estimation method that uses local instrumental or generalized method-of-moments techniques. We leave this as a desideratum for future research.

References

Asaftei, Gabriel, Christopher F. Parmeter, and Yuan Yuan. 2008. "Economies of Scope in Financial Services: A Semi-Parametric Approach." Virginia Tech Working Paper, No. 2008-06.

Berg, Sigbjorn A., and Moshe Kim. 1998. "Banks as Multioutput Oligopolies: An Empirical Evaluation of the Retail and Corporate Banking Markets." *Journal of Money, Credit, and Banking* 30 (2): 135-153.

Blundell, Richard, and Monica C. Dias. 2008. "Alternative Approaches to Evaluation in Empirical Microeconomics." cemmap working paper, No. CWP26/08.

Borenstein, Severin, James B. Bushnell, and Frank A. Wolak. 2002. "Measuring Market Inefficiencies in California's Restructured Wholesale Electricity Market." *American Economic Review* 92 (5): 1376-1405.

Bresnahan, Timothy F. 1989. Empirical Studies of Industries with Market Power. In *Handbook of Industrial Organization*, vol. 2, edited by Richard Schmalensee and Robert D. Willig, 1011-1057. Amsterdam: Elsevier.

Cai, Zongwu, Jianqing Fan, and Runze Li. 2000. "Efficient Estimation and Inferences for Varying Coefficient Models." *Journal of the American Statistical Association* 95 (451): 888-902.

Fabrizio, Kira, Nancy Rose, and Catherine Wolfram. 2007. "Do Markets Reduce Costs? Assessing the Impact of Regulatory Restructuring on U.S. Electric Generation Efficiency." *American Economic Review* 97 (4), 1250-77.

Fan, Jianqing. 1992. "Design-Adaptive Nonparametric Regression." *Journal of the American Statistical Association* 87 (420): 998-1004.

Fan, Jianqing, and Wenyang Zhang. 1999. "Statistical Estimation in Varying-Coefficient Models." *Annals of Statistics* 27 (5): 1491-1518.

Frolich, Markus. 2008. "Parametric and Nonparametric Regression in the Presence of Endogenous Control Variables." *International Statistical Review* 76 (2): 214-27.

Genesove, David, and Wallace P. Mullin. 1998. "Testing Static Oligopoly Models: Conduct and Cost in the Sugar Industry, 1890-1914." Rand Journal of Economics 29 (2): 355-77.

Hall, Robert E. 1988. "The Relation between Price and Marginal Cost in U.S. Industry." *Journal of Political Economy* 96 (5): 921-47.

Hartarska, Valentina, Christopher F. Parmeter, and Denis Nadolnyak. 2010. "Economies of Scope of Lending and Mobilizing Deposits in Microfinance Institutions: A Semiparametric Analysis." *American Journal of Agricultural Economics* 93 (2): 389-98.

Hoover, Donald R., John A. Rice, Colin O. Wu and Li-Ping Yang. 1998. "Nonparametric Smoothing Estimates of Time-Varying Coefficient Models with Longitudinal Data." *Biometrika* 85 (4): 809-22.

Kim, Dae-Wook, and Christopher R. Knittel. 2006. "Biases in StaticOligopoly Models? Evidence from the California Electricity Market." *Journal of Industrial Economics* 54 (4): 451-70.

Koetter, Michael, James W. Kolari, and Laura Spierdijk. 2011. "Enjoying the Quite Life under Deregulation? Evidence from Unadjusted Lerner Indices for U.S. Banking." *Review of Economics and Statistics*, forthcoming.

Konings, Jozef, Patrick Van Cayseele, and Frederic Warzynski. 2005. "The Effects of Privatizations and Competitive Pressure on Firms' Price-Cost Margins: Evidence from Emerging Economies." *Review of Economics and Statistics* 87 (1): 124-34.

Kumbhakar, Subal C, and Knox C.A. Lovell. 2000. *Stochastic Frontier Analysis*. Cambridge: Cambridge University Press.

Kumbhakar, Subal C., Byeong U. Park, Leopold Simar, and Efthymios G. Tsionas. 2007. "Nonparametric Stochastic Frontiers: A Local Maximum Likelihood Approach." *Journal of Econometrics* 137 (1): 1-27.

Kumbhakar, Subal C., and Efthymios G. Tsionas. 2008. "Scale and Efficiency Measurement using a Semiparametric Stochastic Frontier Model: Evidence from the U.S. Commercial Banks." *Empirical Economics* 34 (3): 585-602.

Li, Qi, Cliff J. Huang, Dong Li, Tsu-Tan Fu. "Semiparametric Smooth Coefficient Models." *Journal of Business and Economic Statistics* 20 (3): 412-22.

Mamuneas, Theofanis P., Andreas Savvides, and Thanasis Stengos. 2006. "Economic Development and the Return to Human Capital: A Smooth Coefficient Semiparametric Approach." *Journal of Applied Econometrics* 21 (1): 111-32.

Nevo, Aviv. 2001. "Measuring Market Power in the Ready-to-Eat Cereal Industry." *Econometrica*, 69 (2): 307-42.

Reiss, Peter C., and Frank A. Wolak. 2007. "Structural Econometric Modeling: Rationales and Examples from Industrial Organization." In *Handbook of Econometrics*, vol. 2, edited by James J. Heckman and Edward E. Leamer, 4277-4415. Amsterdam: Elsevier.

Roeger, Werner. 1995. "Can Imperfect Competition Explain the Difference between Primal and Dual Productivity Measures? Estimates for U.S. Manufacturing." *Journal of Political Economy* 103 (2), 316-30.

Rosse, James N. 1967. "Daily Newspapers, Monopolistic Competition, and Economies of Scale." *American Economic Review* 57 (2): 522-33.

Stone, Charles J. 1977. "Consistent Nonparametric Regression." *Annals of Statistics* 5 (4): 595-620.

Wolfram, Catherine D. 1999. "Measuring Duopoly Power in the British Electricity Spot Market." *American Economic Review* 89 (4): 805-26.

Zhang, Wenyang, Sik-Yum Lee, Xinyuan Song. 2002. "Local Polynomial Fitting in Semivarying Coefficient Model." *Journal of Multivariate Analysis* 82 (1): 166-88.

TABLE 1 – SUMMARY STATISTICS OF ACTUAL DATA

Variable	Mean	Std. dev.	Min.	Max.
Panel A. Actual data				
mc	34.45	17.97	18.68	441.11
png	27.30	9.02	15.65	47.42
pno	3.87	8.16	0.13	36.59
q	5,006.8	3,215.3	539.0	15,632.5
tc	181,056.7	179,676.7	12,914.9	1,067,486.0
Panel B. Simulated panel data				
mc	34.40	4.90	20.28	61.23
png	27.24	7.40	6.42	65.17
pno	3.85	2.78	0.01	32.93
q	5,007.1	99.8	4,599.5	5,390.3
tc	218,136.7	36,870.9	143,711.3	328,261.3

Notes: The table reports summary statistics for the actual data (Panel A) and the simulated data (Panel B). The mc is the true marginal cost, png is the average daily price of natural gas for California, pno is the price of NO_x permits, and q is the quantity of output produced and tc is the total cost. The actual data are hourly time series (number of observations is 21,217). The simulated panel data include information for 4,000 time series over 5 firms (20,000 observations).

TABLE 2 – ESTIMATION OF PARAMETRIC MODELS USING ACTUAL DATA

	(1)	(2)	(3)	(4)
Functional form:	log-linear	translog	log-linear	translog
Estimation method:	OLS	OLS	2SLS	2SLS
ln(png/pno)	1.016	0.896	1.009	-0.111
	(926.31)	(51.11)	(892.36)	(-2.57)
lnq	1.105	0.538	1.069	-0.797
	(547.17)	(13.55)	(448.99)	(-10.87)
½*[ln(png/pno)]^2		0.084		0.130
		(52.19)		(52.21)
½*(lnq)^2		0.071		0.188
		(15.37)		(24.34)
ln(png/pno)*lnq		-0.013		0.089
		(-7.39)		(20.58)
Constant	-0.788	1.669	-0.459	8.996
	(-42.82)	(9.61)	(-21.30)	(25.17)
Observations	21,217	21,217	21,217	21,217
Adj. R-squared	0.977	0.981	0.977	0.978

Notes: The table reports the results (coefficients and *t*-statistics) from the estimation of Equations (2) and (3) using th,e actual data from Borenstein et al. (2002). All coefficient estimates are statistically significant at the 1% level. The dependent variable is the natural logarithm of total cost divided by the price of NOx permits, i.e. ln(tc/pno). The png is the average daily price of natural gas for California, and q is the quantity of output produced at each point in time.

TABLE 3 – SUMMARY STATISTICS OF ESTIMATES OF MARGINAL COST FROM PARAMETRIC MODELS USING ACTUAL DATA

	Mean	Std. dev.	Min.	Max.
mc_{llols}	34.71	12.78	20.88	81.17
mc_{trols}	34.68	14.60	18.25	88.59
mc_{ll2sls}	35.27	12.99	21.22	82.48
mc_{tr2sls}	32.82	11.70	14.14	80.18

Notes: The table reports summary statistics for the estimates of marginal cost obtained from the four alternative parametric models of Table 2. The mc_{llols} is from the estimation of the log-linear specification with OLS, mc_{trolls} is from the estimation of the translog specification with OLS, mc_{ll2sls} is from the estimation of the log-linear specification with a two-stage least squares, and mc_{tr2sls} is from the estimation of a translog specification with a two-stage least squares.

TABLE 4 – CORRELATIONS BETWEEN THE VALUES OF THE TRUE MARGINAL COST AND THE ESTIMATED MARGINAL COST FROM PARAMETRIC MODELS AND ACTUAL DATA

	mc	mc_{llols}	mc_{trols}	mc_{ll2sls}	mc_{tr2sls}
mc	1.000				
mc_{llols}	0.898	1.000			
mc_{trols}	0.902	0.997	1.000		
mc_{ll2sls}	0.898	1.000	0.997	1.000	
mc_{tr2sls}	0.879	0.936	0.954	0.936	1.000

Notes: The table reports pairwise correlation coefficients between the true values of marginal cost (mc) and the estimates of marginal cost obtained from the four alternative specifications in Table 2. The mc_{llols} is from the estimation of the log-linear specification with OLS, mc_{trolls} is from the estimation of the translog specification with OLS, mc_{ll2sls} is from the estimation of the log-linear specification with two-stage least squares, and mc_{tr2sls} is from the estimation of a translog specification with a two-stage least squares.

TABLE 5 – ESTIMATION OF PARAMETRIC MODELS USING SIMULATED DATA

	(1)	(2)
Functional form:	log-linear	translog
Estimation method:	OLS	OLS
ln(png/pno)	0.898***	-2.133**
	(371.11)	(-2.07)
lnq	0.941***	41.048
	(8.99)	(0.65)
½*[ln(png/pno)]^2		0.074***
		(21.60)
½*(lnq)^2		-4.795
		(-0.65)
ln(png/pno)*lnq		0.334***
		(2.77)
Constant	1.219	-166.26
	(1.37)	(-0.62)
Observations	20,000	20,000
Adj. R-squared	0.862	0.865

Notes: The table reports the results (coefficients and *t*-statistics) from the estimation of Equations (2) and (3) using the simulated data. The dependent variable is the natural logarithm of total co,st divided by the price of Nox permits, i.e. ln(tc/pno).The png is the average daily price of natural gas for California, and q is the quantity of output produced at each point in time. The *** and ** denote statistical significance at the 1 and 5% levels respectively.

TABLE 6 – SUMMARY STATISTICS OF ESTIMATES OF MARGINAL COST FROM PARAMETRIC MODELS USING SIMULATED DATA

	Mean	Std. dev.	Min.	Max.
mc_{llols}	41.70	7.02	28.53	62.33
mc_{trols}	42.37	14.12	6.64	139.81

Notes: The table reports summary statistics for the estimates of marginal cost obtained from the two alternative parametric models of Table 5. The mc_{llols} is from the estimation of the log-linear specification with OLS, and mc_{trolls} is from the estimation of the translog specification with OLS.

TABLE 7 – CORRELATIONS BETWEEN THE VALUES OF THE TRUE MARGINAL COST AND THE ESTIMATED MARGINAL COST FROM PARAMETRIC MODELS AND SIMULATED DATA

	mc	mc_{llols}	mc_{trols}
mc	1.000		
mc_{llols}	-0.005	1.000	
mc_{trols}	0.342	0.105	1.000

Notes: The table reports pairwise correlation coefficients between the true values of marginal cost (mc) and the estimates of marginal cost obtained from the two alternative specifications in Table 5. The mc_{llols} is from the estimation of the log-linear specification with OLS, and mc_{trolls} is from the estimation of the translog specification with OLS.

TABLE 8 – SUMMARY STATISTICS OF THE ESTIMATED COEFFICIENTS FROM THE SMOOTH COEFFICIENT MODEL

Panel A. Actual data

	Mean	Std. dev.	Min.	Max.
png	27.30	9.02	15.65	47.42
pno	11.70	0.90	9.47	13.88
q (mc_{scm})	35.10	5.86	31.24	56.70
Correlation coefficient between mc_{scm} and mc = 0.974.				

Panel B. Simulated data

	Mean	Std. dev.	Min.	Max.
png	27.24	7.40	6.42	65.17
pno	27.24	7.40	6.42	65.17
q (mc_{scm})	34.70	2.47	31.17	50.68
Correlation coefficient between mc_{scm} and mc = 0.950.				

Notes: The table reports summary statistics for the estimates from the smooth coefficient model. Panel A shows the results from the data set of Borenstein et al. (2002), and Panel B shows the results from the simulated data. The png is the average daily price of natural gas for California, pno is the price of Nox permits, and q is the quantity of output produced at each point in time that also equals the marginal cost from the smooth coefficient model (mc_{scm}).

TABLE 9 – RELATION BETWEEN THE ESTIMATED AND THE TRUE MARGINAL COST

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimated mc	1.263	1.115	1.243	1.349	-0.004	0.027	1.036	1.137
	(297.97)	(312.38)	(297.97)	(267.81)	(-0.77)	(10.85)	(407.69)	(432.18)
Constant	-9.372	-4.206	-9.372	-9.808	-34.562	-35.529	-2.151	-11.448
	(-59.79)	(-31.32)	(-59.79)	(-55.90)	(-165.58)	(-325.17)	(-4.62)	(-17.35)
R-squared	0.807	0.821	0.807	0.772	0.006	0.160	0.969	0.931

Notes: The table reports coefficient estimates of the regression y=a+bx+u where y is the true marginal cost and x is the estimated marginal cost from the eight different specifications presented in Tables 2, 5, and 8. Specifications (1) to (4) use the marginal cost values obtained from the four equations presented in Table 2. Specifications (5) and (6) use the marginal cost values obtained from the two equations presented in Table 5. Specifications (7) and (8) use the marginal cost values obtained from the two equations presented in Panels A and B of Table 8.

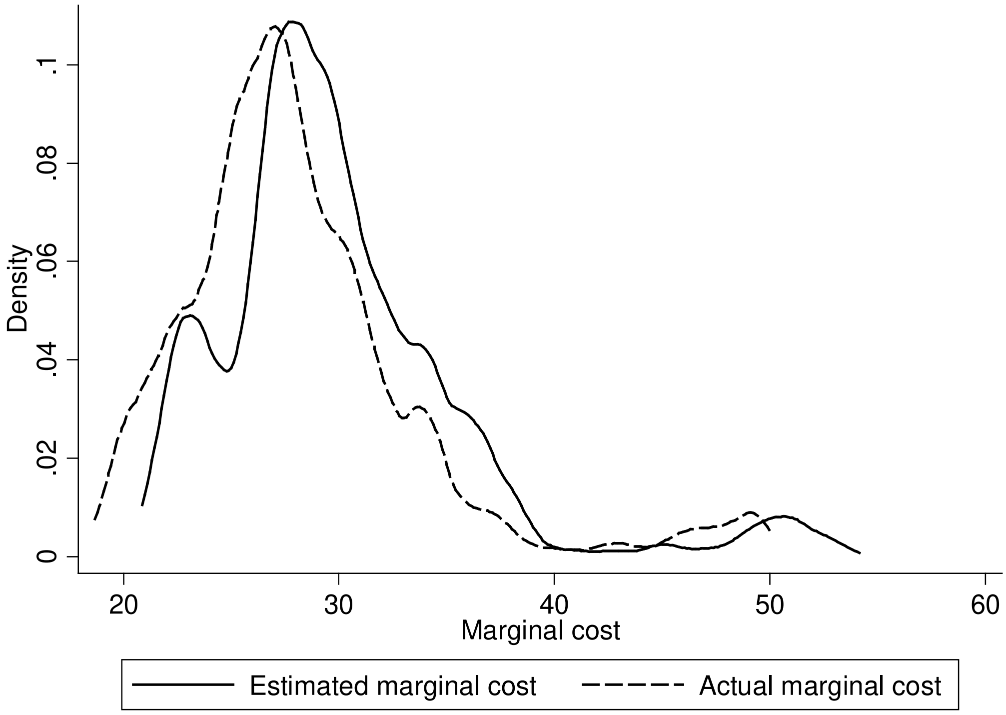

FIGURE 1. MARGINAL COST OBTAINED FROM THE LOG-LINEAR SPECIFICATION AND OLS VS. TRUE MARGINAL COST (ACTUAL DATA)

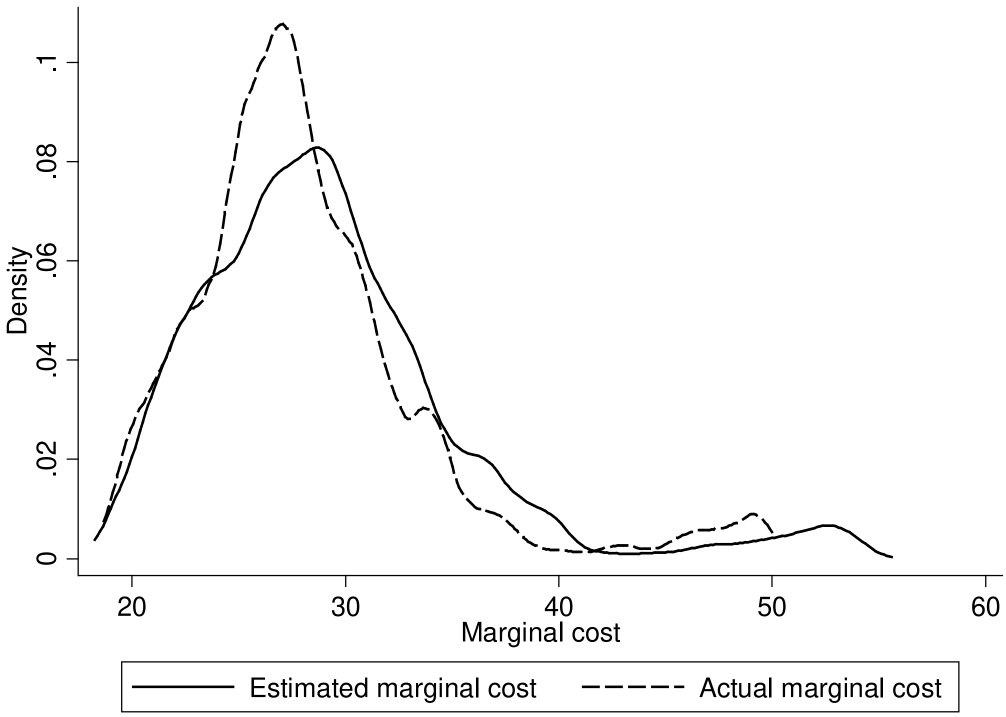

FIGURE 2. MARGINAL COST OBTAINED FROM THE TRANLOG SPECIFICATION AND OLS VS. TRUE MARGINAL COST (ACTUAL DATA)

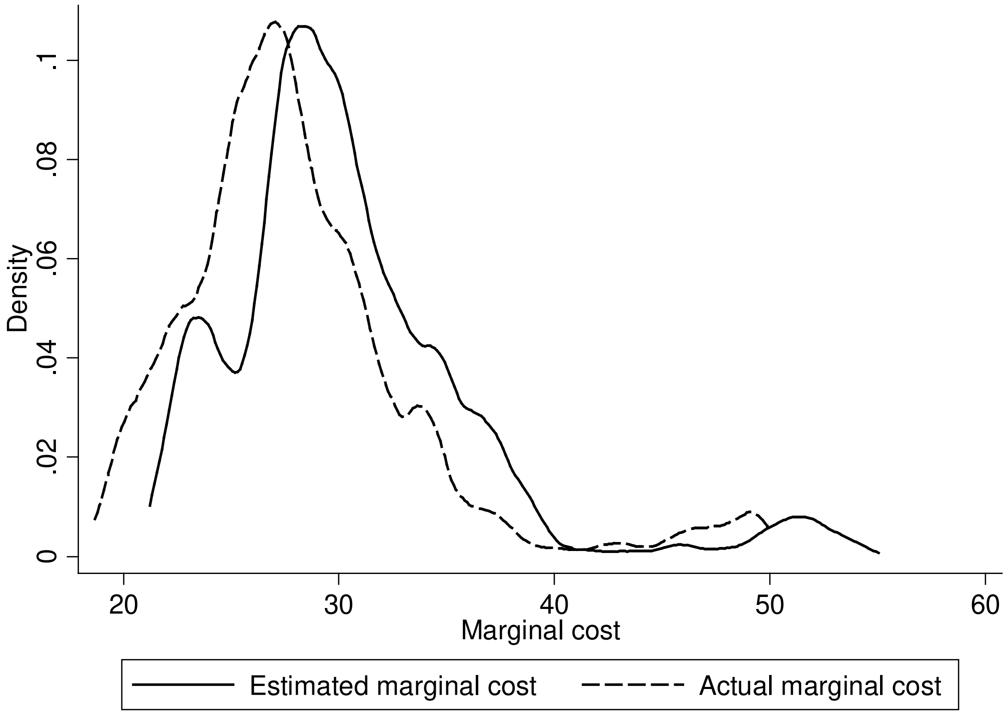

FIGURE 3. MARGINAL COST OBTAINED FROM A LOG-LINEAR SPECIFICATION AND 2SLS VS. TRUE MARGINAL COST (ACTUAL DATA)

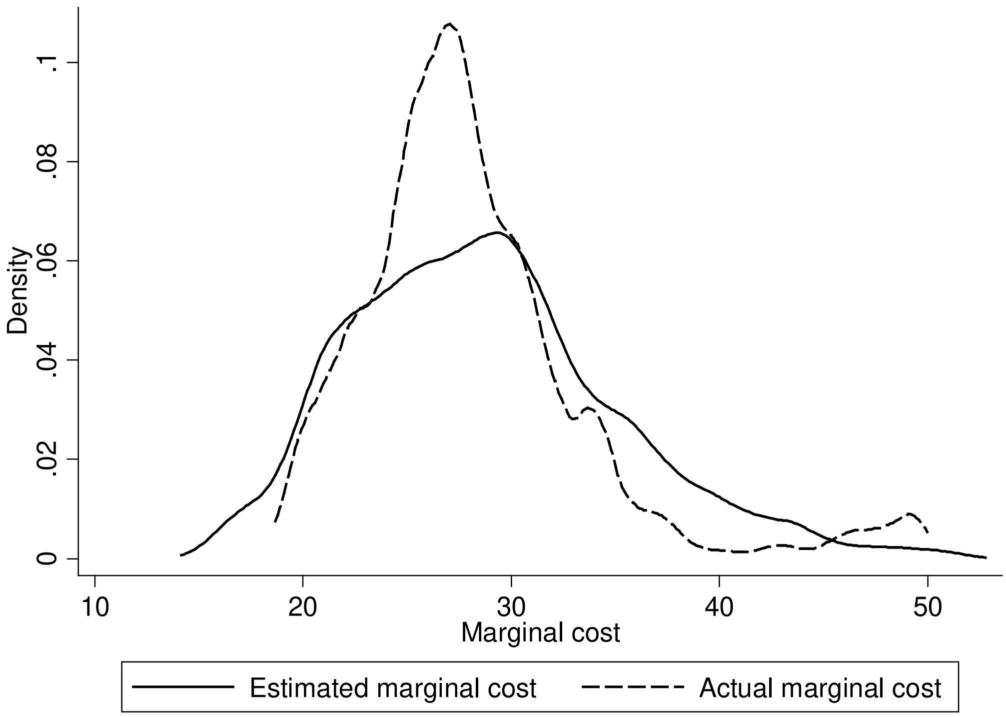

FIGURE 4. MARGINAL COST OBTAINED FROM A TRANSLOG SPECIFICATION AND 2SLS VS. TRUE MARGINAL COST (ACTUAL DATA)

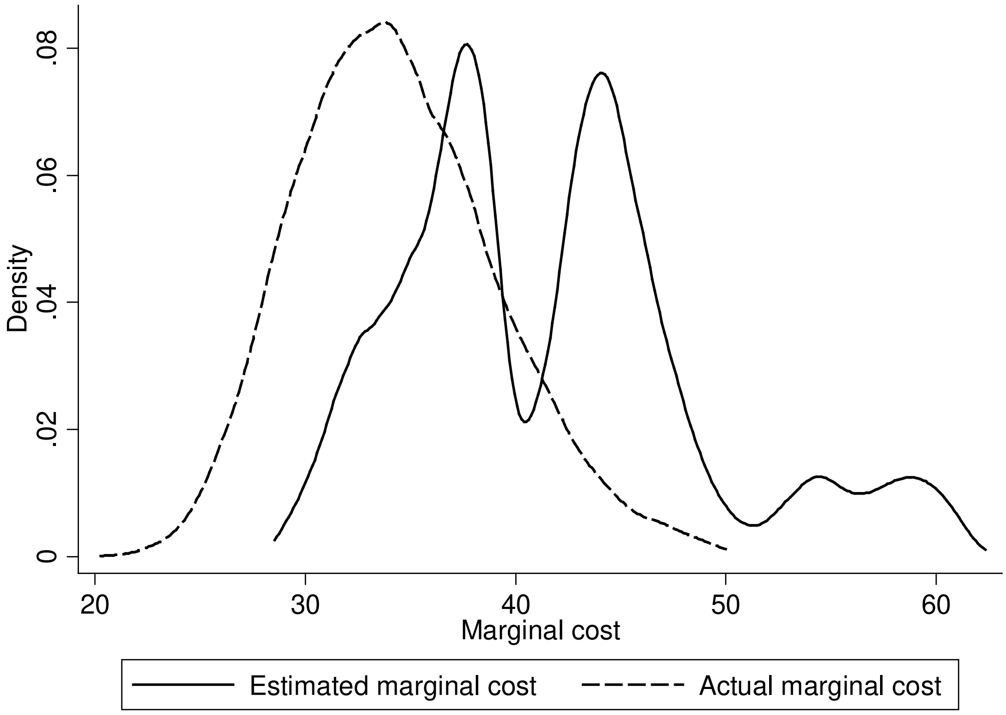

FIGURE 5. MARGINAL COST OBTAINED FROM THE LOG-LINEAR SPECIFICATION AND OLS VS. TRUE MARGINAL COST (SIMULATED DATA)

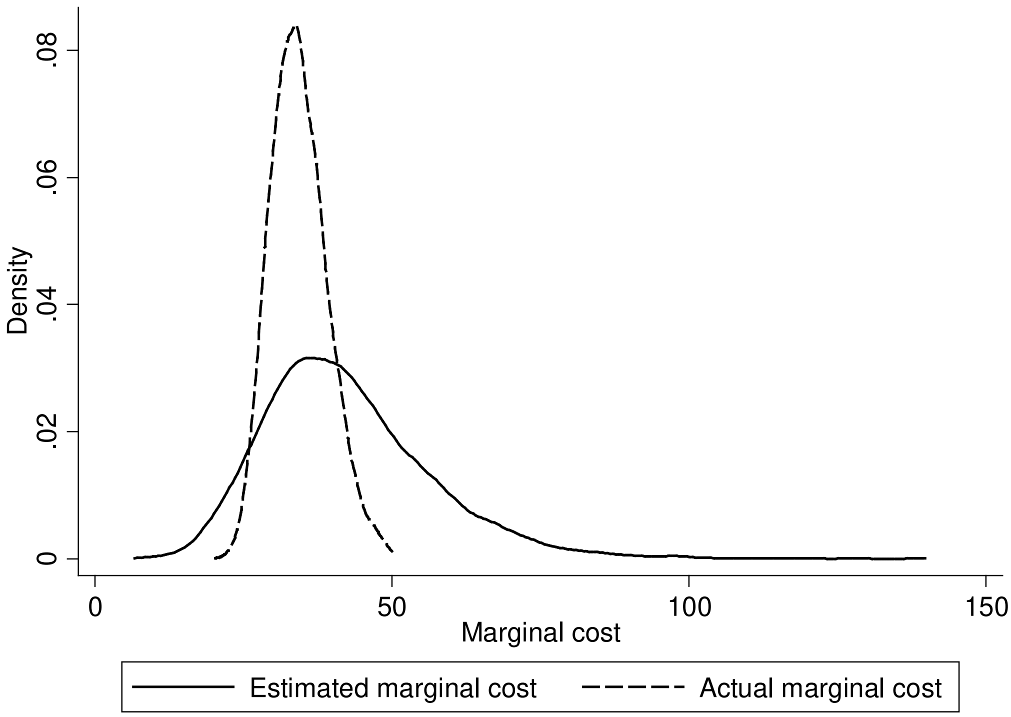

FIGURE 6. MARGINAL COST OBTAINED FROM THE TRANSLOG SPECIFICATION AND OLS VS. TRUE MARGINAL COST (SIMULATED DATA)

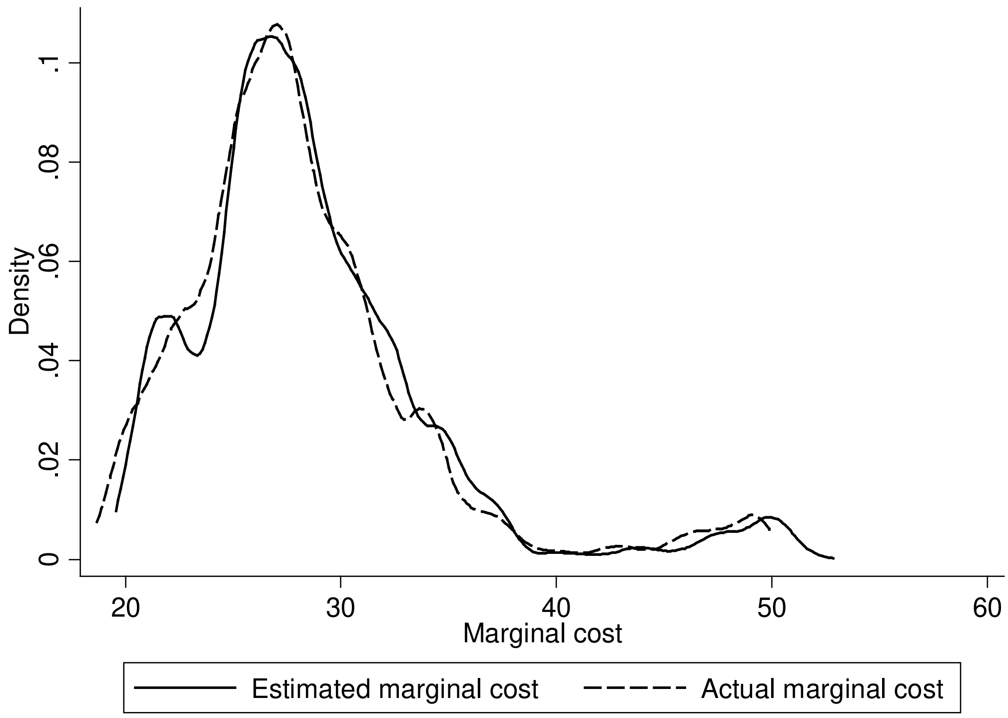

FIGURE 7. MARGINAL COST OBTAINED FROM THE SMOOTH COEFFICIENT MODEL VS. TRUE MARGINAL COST (ACTUAL DATA)

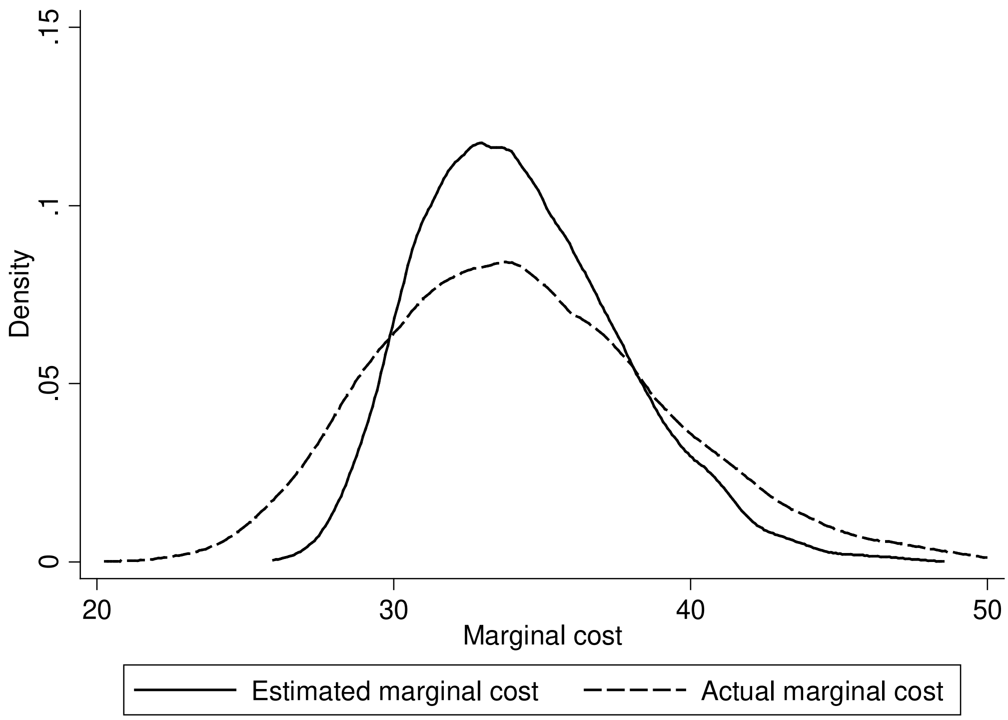

FIGURE 8. MARGINAL COST OBTAINED FROM THE SMOOTH COEFFICIENT MODEL VS. TRUE MARGINAL COST (SIMULATED DATA)