CrossMark

Estimation of cost efficiency without cost data

Levent Kutlu [1] · Ran Wang[2]

Abstract
One of the advantages of conduct parameter games is that they enable estimation of market power without total cost data. In line with this, we develop a conduct parameter based model to estimate the firm specific "marginal cost efficiency" and conduct without using total cost data. The marginal cost efficiency is an alternative measure of efficiency that is based on deadweight loss. We illustrate our methodology by estimating firm-route-quarter specific conducts and marginal cost efficiencies of U.S. airlines for Chicago based routes without using route-level total cost data.

Keywords Market power · Conduct parameter · Efficiency · Stochastic frontier analysis · Airlines

JEL classification C13 · L13 · L93

1 Introduction

The Lerner index Lerner (1934) is a widely used market power measure, which is the ratio of price-marginal cost mark-up and price. One of the difficulties for calculating the Lerner index is that total cost data may not be available making the estimation of the marginal cost difficult. A solution to this problem is estimating a conduct parameter (or conjectural variations) game[1] in which firms form conjectures about the variation in other firms' strategies (e.g., output) in response to a change in their own strategies. For given demand and cost conditions, the conjectures corresponding to the observed price-cost margins can be estimated "as-if" the firms are playing a conduct parameter game. In this setting, the "implied marginal cost" can be estimated via a supply-demand system.

Stochastic frontier analysis (SFA) literature[2] suffers from a similar problem. That is, the standard SFA models require

total cost data in order to estimate the cost efficiency of a firm. Moreover, as we will describe in the next section, market power and efficiency are closely related concepts and ignoring inefficiency in a conduct parameter model may lead to inconsistent conduct parameter estimates. What is more, ignoring the inefficiencies of productive units may invalidate the standard deadweight loss (DWL) calculations since DWL from collusive behavior depends on inefficiency levels.[3] If the productive units exhibit inefficiency that is misinterpreted as firm heterogeneity, then the standard calculations of DWL become invalid. In such cases, Kutlu and Sickles (2012) recommend using what they call the efficient full marginal cost (EFMC) for markup calculation.[4] Hence, estimation of conduct, marginal cost, and marginal cost efficiency would be essential for a valid DWL

✉ Levent Kutlu
  levent.kutlu@econ.gatech.edu

[1] School of Economics, Georgia Institute of Technology, Atlanta, GA, USA

[2] SunTrust Bank, 303 Peachtree St. NE, Atlanta, GA, USA

[1] See Perloff et al. (2007) and Bresnahan (1989) for more details on conduct parameter approach.

[2] See Kumbhakar and Lovell (2003) for a book-length survey on SFA and Sickles (2005) for a simulation study examining the performances of some estimators in the SFA literature.

[3] See Comanor and Leibenstein (1969) and Kutlu and Sickles (2012) for more details about calculation of DWL when firms are inefficient.

[4] Kutlu and Sickles (2012) define EFMC as the sum of a shadow cost and efficient marginal cost that is calculated using stochastic frontier analysis techniques. The shadow cost reflects the constraints that the firms face such as capacity or incentive compatibility constraints. Efficient marginal cost is the marginal cost when the firm achieves full efficiency. Although they define the EFMC concept, they calculate EFMC using cost efficiency estimates obtained from a standard stochastic cost frontier model. Hence, while the concept is due to Kutlu and Sickles (2012), arguably, we develop a more proper method that can directly calculate EFMC.

calculation. We overcome all these issues by combining conduct parameter and SFA literatures so as to generalize a conventional conduct parameter model to allow inefficiency in marginal cost. This enables the estimation of marginal cost efficiencies and conduct parameters jointly and consistently without using total cost data. In contrast to the SFA literature, which infers cost efficiency from a cost function, we estimate the marginal cost efficiency from a supply-demand system that is derived from our conduct parameter game. This introduces a related but different measure of efficiency that is based on deadweight loss (DWL), i.e., marginal cost efficiency. Therefore, the marginal cost efficiency concept would be a valuable tool for antitrust authorities.

We would like to note that our methodology can be applied to a variety of existing conduct parameter settings. Hence, we consider our study as a guideline for estimating conduct parameter models in the presence of (marginal cost) inefficiency. Examples for how this can be accomplished are illustrated in the Appendix. Hence, our study provides a link between industrial organization and SFA literatures to enable researchers to model inefficiency in more structural settings.

In order to illustrate how our methodology can be used in an empirical framework, we apply our methodology to estimate the firm-route-quarter specific conducts and marginal cost efficiencies of U.S. airlines for routes that originate in Chicago.[5] The time period that our data set covers is 1999I–2009IV. One of the difficulties that empirical researchers face is that the availability of cost data. Aggregate airline cost data is available but not at the route level. Therefore, route level total cost data is not available. Kutlu and Sickles (2012) try to overcome this problem by incorporating a specific number of enplanements for each airline, a specific distance of each city-pair, and airline fixed effects when estimating the cost function.[6] This enables the calculation of firm-route-quarter specific marginal costs from the cost function estimation. However, their efficiency estimates are still aggregate firm-quarter specific making it difficult to understand route-specific difficulties that may increase inefficiency.[7] Moreover, Kutlu and Sickles (2012) first estimate cost efficiencies using a standard stochastic frontier model.[8] These derived efficiency estimates are used

to estimate the supply relation.[9] The conduct estimates are conditional on these derived efficiency estimates. In contrast to their study, we jointly estimate the firm-route-quarter specific conducts and marginal cost efficiencies of the U.S. airlines without the use of route specific total cost data.[10] Compared to the standard stochastic frontier efficiency measures, our efficiency concept is a more relevant measure from the antitrust point of view. Moreover, our empirical results for the relationship between market concentration and efficiency can be used as a robustness check from another perspective. Our results suggest that concentration ratio (measured by CR_4) and market share are negatively related to the marginal cost efficiency. In contrast to this, the concentration ratio and market share are positively related to the conduct. Finally, we find that both the conduct and DWL estimates may be biased if a conventional conduct parameter model, which ignores marginal cost inefficiency, is used.

The rest of the paper is structured as follows. In Section 2, we briefly discuss the relationship between market power and efficiency. In Section 3, we build up our theoretical model. In Section 4, we describe our data set, present our empirical model, and discuss our results. In the next section, we make our concluding remarks. Finally, in the Appendix, we present extensions of the theoretical model.

2 Market power and efficiency

"The Quiet Life Hypothesis" (QLH) by Hicks (1935) and "the Efficient Structure Hypothesis" (ESH) by Demsetz (1973) are two well-known hypotheses that relate market power to efficiency. The former claims that higher competitive pressure is likely to force management work harder, which in turn increases efficiencies of firms. The latter states that the firms with superior efficiency levels use their competitive advantages to gain larger market shares, which leads to higher market concentration and thus higher market power. The findings of Berger and Hannan (1998) and Kutlu and Sickles (2012) support the QLH for the banking and airline industries, respectively. However, Maudos and Fernandez de Guevara (2007) show evidence for ESH applicable to the banking industry. Delis and Tsionas (2009) are in favor of the QLH on average but acknowledge that the ESH may prevail in the case of highly efficient banks. The relationship between market power and efficiency has long been acknowledged by economists.

---

[5] See Berry and Jia (2010) for a paper that is studying airline performance in a different framework.

[6] Weiher et al. (2002) use a similar approach.

[7] Since Kutlu and Sickles (2012) estimate an aggregate model (i.e., route specific market power), they use market share weighted efficiency estimates in their estimations. That is why their route specific efficiency variables are not the same for different routes.

[8] Kutlu and Sickles (2012) use a variation of the distribution free approach proposed by Cornwell et al. (1990). Kutlu (2017b) discusses some of the potential estimation difficulties for such distribution free approaches.

[9] When calculating DWL values, Kutlu and Sickles (2012) used SFA-type efficiency estimates as a proxy for marginal cost efficiency.

[10] Similar to our study, Delis and Tsionas (2009) simultaneously estimate bank conducts and efficiencies. However, their model requires total cost data. Hence, as it stands, their methodology is not applicable to our airline example.

However, market power and SFA literatures largely ignore this relationship. Delis and Tsionas (2009), Koetter and Poghosyan (2009), Koetter et al. (2012), and Kutlu and Sickles (2012) exemplify some studies that attempt to estimate the market powers of firms in frameworks where firms are allowed to be inefficient. Except for Delis and Tsionas (2009), the market power estimates in these studies are conditional on efficiency estimates.

In conduct parameter models, ignoring marginal cost inefficiency can potentially cause inconsistent conduct parameter estimates. In particular, the conduct parameter estimates may pick up some of the marginal cost inefficiency. Delis and Tsionas (2009) argue that if the inefficiency is not taken into account, the optimization model of firms become irrelevant, which would lead to severe bias as the level of inefficiency increases. Similarly, ignoring market power (conduct parameter) can potentially cause inconsistent cost efficiency or marginal cost efficiency estimates. The differences in market powers would lead to different firm behavior and this can be confused with the firm level inefficiency. Generally, efficiencies are measured by closeness of production units to the best-practice units observed in the market. If the firm level conducts affect the performance of the best-practice units, then the efficiency estimates which do not take this into account would not be accurate. For instance, in a market facing a Cournot competition the best practicing firm may not really be fully efficient.[11] As mentioned in the introduction, our methodology aims to overcome these difficulties by explicitly and simultaneously modeling conduct parameter and marginal cost efficiency.

3 Theoretical model

In this section, we describe the theoretical framework, used to estimate marginal cost efficiencies and conducts of firms without total cost data. The stochastic frontier literature relaxes full efficiency assumption of neoclassical production theory by allowing the firms to be inefficient. The inefficiency is treated as an unobserved component, which is captured by a one-sided error term. In the conventional stochastic frontier framework, the cost efficiencies of firms would be estimated by the following model:

$$\ln C\left(q_{it}; X_{c,it}\right) = \ln C^*\left(q_{it}; X_{c,it}\right) + u_{it} + v_{it}. \tag{1}$$

where q_{it} is the quantity of firm i at time t; $X_{c,it}$ is a vector of variables related to cost; $u_{it} \geq 0$ is a term which is capturing the inefficiency; v_{it} is the usual two-sided error term; and C^*

is the deterministic component of cost when firms achieve full efficiency. A variety of distributions are proposed for u_{it} including the half normal (Aigner et al. 1977), exponential (Meeusen and van den Broeck 1977), truncated normal (Stevenson 1980), and gamma (Greene 1980a, 1980b, 2003) distributions. The cost efficiency of a firm, EFF_{it}, is estimated by:

$$\begin{aligned} EFF_{it} &= \exp(-\hat{u}_{it}) \\ \hat{u}_{it} &= E[u_{it}|u_{it} + v_{it}]. \end{aligned} \tag{2}$$

The stochastic frontier approach requires detailed cost data, which many times is not available. We utilize the conduct parameter approach to overcome this issue. For this purpose, instead of modelling total cost as in the conventional SFA models, we directly model marginal cost, c, as follows:

$$\ln c\left(q_{it}; X_{c,it}\right) = \ln c^*\left(q_{it}; X_{c,it}\right) + u_{it} + v_{it}. \tag{3}$$

where c^* is the deterministic component of marginal cost when firms achieve full efficiency; $u_{it} \geq 0$ is a term which is capturing the marginal cost inefficiency; and v_{it} is a two-sided random variable, which is observed by the firm but not observed by the researcher. We call c^* efficient marginal cost (EMC). Rather than estimating a cost function, we estimate a supply-demand system that enables us to calculate the marginal cost efficiency. From the antitrust point of view, which is concerned with market power and DWL estimations, the marginal cost efficiency is a more relevant efficiency concept compared to the cost efficiency concept.

Let $P_t = P(Q_t; X_{d,t})$ be the inverse demand function, Q_t be the total quantity, and $X_{d,t}$ is a vector of demand related variables at time t. The perceived marginal revenue (PMR) is given by:

$$\begin{aligned} PMR(\theta_{it}) &= P_t + \frac{\partial P_t}{\partial Q_t}\frac{\partial Q_t}{\partial q_{it}}q_{it} \\ &= P_t\left(1 - \frac{s_{it}}{E_t}\theta_{it}\right) \end{aligned} \tag{4}$$

where $s_{it} = \frac{q_{it}}{Q_t}$ is the market share of firm i at time t; $E_t = -\frac{\partial Q_t}{\partial P_t}\frac{P_t}{Q_t}$ is the (absolute value of) elasticity of demand; $\theta_{it} = \frac{\partial Q_t}{\partial q_{it}}$ is the conduct parameter. Three benchmark values for $\theta_{it} = \left\{0, 1, \frac{1}{s_{it}}\right\}$ correspond to perfect competition, Cournot competition, and joint profit maximization, respectively. The supply relation is:[12]

$$\begin{aligned} &P_t\left(1 - \frac{s_{it}}{E_t}\theta_{it}\right) = c_{it} \Leftrightarrow \\ &\ln P_t + \ln\left(1 - \frac{s_{it}}{E_t}\theta_{it}\right) = \ln c_{it} \end{aligned} \tag{5}$$

---

[11] Lee and Johnson (2015) argue that in imperfectly competitive markets inefficiency may in fact be a result of endogenous prices and the effect of output production on price.

[12] Note that perceived marginal revenue must be positive so that the equilibrium makes sense. Hence, we assume that $1 - \frac{s_{it}}{E_t}\theta_{it} > 0$. So, we have $\ln\left(1 - \frac{s_{it}}{E_t}\theta_{it}\right) \leq 0$.

where $c_{it} = c(q_{it}; X_{c,it})$. After including the econometric error terms, the supply relation becomes:[13]

$$\ln P_{it} = \ln c_{it}^* + g_{it} + u_{it} + \varepsilon_{it}^s \qquad (6)$$

where $g_{it} = g(\theta_{it}, s_{it}, E_t) = -\ln\left(1 - \frac{s_{it}}{E_t}\theta_{it}\right) \geq 0$ is the market power term, which is an increasing function of θ_{it}; $u_{it} \geq 0$ is the inefficiency term; and ε_{it}^s is the two-sided error term.

The E_t term is identified through the demand equation. Intuitively, Eq. (4) suggests that if c_{it} and q_{it} are highly collinear, then the conduct parameter may be identified through the variation in $\frac{\partial P_t}{\partial Q_t}$. We assume that the demand and marginal cost functions are such that the conduct parameter and marginal cost can be separately identified.[14] In most cases, identification is a problem when the demand function is linear. For example, when the demand and marginal cost functions are linear, we do not observe a variation in $\frac{\partial P_t}{\partial Q_t}$, and c_{it} and q_{it} are perfectly collinear. In this case, we cannot separately identify the conduct parameter and marginal cost. One way to achieve identification is assuming constant marginal cost.[15] Another commonly used approach that does not require constant marginal cost assumption is including the cross-product of quantity and an exogenous variable in the demand equation. When such cross-products are included in the model, the identification of conduct parameter is achieved through both parallel shifts and rotations of the demand curve. Bresnahan (1982) illustrates how identification can be achieved by such rotations for the linear demand and linear marginal cost case. He states that the logic of identification is maintained even if the curves are not linear. In general, the conduct parameter is identifiable if the inverse demand function is not separable in exogenous variables, Z, and the number of exogenous variables is enough. More precisely, we can write the inverse demand function P so that $P = f(Q, r(Z))$ where Z is a vector of exogenous variables but P does not take the form $P = Q^{-1/\theta}r(Z) + h(Q)$ for some functions f, r, and h if and only if the identification is impossible.

Our model is different from the standard market power models due to the additional u_{it} term. This inefficiency term, u_{it}, is identified by utilizing the asymmetric distribution of the composed error term, i.e., $u_{it} + \varepsilon_{it}^s$. Intuitively, u_{it} is identified if the signal-to-noise ratio (the variance ratio of the inefficiency component to the composed error) is not small. Hence, the identification of model parameters
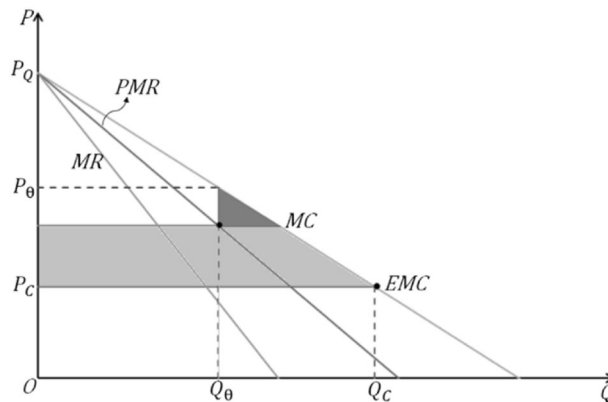

Fig. 1 Conduct, marginal cost efficiency, and social welfare

requires the standard conduct parameter model and stochastic frontier model identification assumptions to hold when there are endogenous variables.[16]

Following Kutlu and Sickles (2012), Fig. 1 aims to illustrate the underlying mechanism of our model and consequences of ignoring inefficiency when calculating DWL. The figure includes inverse demand function, perceived marginal revenue (PMR), marginal revenue (MR) that is corresponding to monopoly scenario, marginal cost (MC), and efficient marginal cost (EMC). For illustrative purposes, we consider the same constant marginal costs, conducts, and efficiencies for each firm. P_θ and Q_θ are the equilibrium price and quantity at conduct level θ. Similarly, P_C and Q_C are price and quantity for the perfect competition scenario in which conduct equals 0. In the figure, it is assumed that under perfect competition there would be no inefficiency. If QLH holds, then as the market power, measured by θ, increases MC diverges from EMC. In our framework, the marginal cost efficiency is defined as EMC/MC. The social welfare loss at conduct level θ would be equal to the shaded area (sum of dark and light shaded regions). In Fig. 1, the efficiency is roughly 60%, which is relatively low.[17] As a result the social welfare loss due to inefficiency is substantial. The conventional DWL value, which is ignoring inefficiency, is given by the dark shaded triangular area; and is much smaller than the overall social welfare loss. When there is heterogeneity either in efficiencies or marginal cost frontiers of firms (under constant MC assumption), the calculation of DWL would be similar except that EMC would be a step function rather than a horizontal line.

Now, we describe how this conduct parameter game would be estimated. We assume that the conduct parameter

---

[13] The introduction of the error term enables us to deviate from a single market price. Also, the price may be considered to be a function of firm specific variables, $X_{d,it}$.

[14] For details about the identification conditions for conduct parameter models, we direct the reader to Bresnahan (1982), Lau (1982), Perloff et al. (2007), and Perloff and Shen (2012).

[15] Note that a constant marginal cost function does not depend on quantity but it may still depend on variables other than the quantity.

[16] At the end of this section, we provide a discussion about identification condition for stochastic frontier models with endogenous regressors.

[17] In the figure, for the sake of illustrating the ideas better, we rather use market level conduct parameter.

θ_{it} is a function of variables, $X_{g,it}$, that affect firm specific market power such as market shares and concentration ratios. Modeling θ_{it} through this function may lead to computational difficulties. In contrast, g_{it} can be modeled directly as a function of $X_{g,it}$ so that θ_{it} is solved after getting the parameter estimates. That is, we can calculate the estimate of θ_{it} as follows:

$$\hat{\theta}_{it} = \frac{\hat{E}_t}{s_{it}}(1 - \exp(-\hat{g}_{it})) \tag{7}$$

where $\hat{E}_t$ and $\hat{g}_{it}$ are the estimates for E_t and g_{it}, respectively. The market power term, g_{it}, is bounded by 0 and $B_{it} = -ln\left(1 - \frac{1}{E_t}\right)$. It follows that the choice of functional form should be so that $g_{it} \in [0, B_{it}]$. In this study, we use:

$$g_{it} = \frac{B_{it}\exp\left(X'_{g,it}\beta_g\right)}{1 + \exp\left(X'_{g,it}\beta_g\right)}. \tag{8}$$

One of the drawbacks of the standard stochastic frontier models is that if the regressors are correlated with v_{it} or u_{it}, then the parameter and efficiency estimates are inconsistent. In this setting, v_{it} and u_{it} terms are assumed to be independent, which can be a questionable assumption in a variety of settings. Similar to Kutlu (2010), Karakaplan and Kutlu (2017a, b), and Kutlu et al. (2018), we use a limited information maximum likelihood based approach to handle the endogeneity issue that occurs when the two-sided error term is correlated with the regressors or u_{it}. The approach solves the endogeneity issue by including a bias correction term in the model. For example, u_{it} can be a function of regressors (e.g., market shares of firms or concentration ratios) that are correlated with the two-sided error term. Consider the following supply relation model with endogenous explanatory variables:

$$
\begin{aligned}
\ln P_{it} &= \ln c^*_{it} + g_{it} + u_{it} + \varepsilon^s_{it} \\
X_{en,it} &= \zeta'_{it}\delta + w_{it} \\
\begin{bmatrix} \tilde{w}_{it} \\ \varepsilon^s_{it} \end{bmatrix} &\equiv \begin{bmatrix} \Sigma_w^{-1/2}w_{it} \\ \varepsilon^s_{it} \end{bmatrix} \sim \mathbf{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} I_m & \rho\sigma_\varepsilon \\ \rho'\sigma_\varepsilon & \sigma^2_\varepsilon \end{bmatrix}\right) \\
u_{it} &= h_{it}\tilde{u}_{it} \\
h_{it} &\geq 0 \\
\tilde{u}_{it} &\sim \mathbf{N}^+\left(\mu_u, \sigma^2_u\right)
\end{aligned} \tag{9}
$$

where P_{it} is the price; $X_{en,it}$ is an $m \times 1$ vector of all endogenous variables used in modelling c^*_{it}, g_{it}, and u_{it}; $\zeta_{it} = I_m \otimes Z_{it}$ where Z_{it} is a $l \times 1$ (with $l \geq m$) vector of all exogenous variables. The irregular term ε^s_{it} is correlated with the regressors but independent of $\tilde{u}_{it}$ conditional on

$X_{en,it}$ and Z_{it}.[18] Hence, ε^s_{it} is independent of u_{it} conditional on $X_{en,it}$ and Z_{it}. Note that ε^s_{it} and u_{it} may still be correlated unconditionally. By applying a Cholesky decomposition of the variance-covariance matrix of $\left[\tilde{w}'_{it} \;\; \varepsilon^s_{it}\right]'$, we can rewrite the supply equation as follows:

$$
\begin{aligned}
\ln P_{it} &= \ln c^*_{it} + g_{it} + \sigma_\varepsilon\rho'\tilde{w}_{it} + u_{it} + \tilde{\varepsilon}^s_{it} \\
&= \ln c^*_{it} + g_{it} + \eta'\left(X_{en,it} - \zeta'_{it}\delta\right) + u_{it} + \tilde{\varepsilon}^s_{it}
\end{aligned} \tag{10}
$$

where $\tilde{\varepsilon}^s_{it} \sim \mathbf{N}\left(0, (1 - \rho'\rho)\sigma^2_\varepsilon\right)$ and $\eta = \sigma_\varepsilon\rho\Sigma_w^{-1/2}$. The parameters of this supply relation can be estimated in one stage with the maximum likelihood estimation method. However, sometimes it is simpler to get the consistent parameter estimates in two stages by first estimating the bias correction term $\eta'\left(X_{en,it} - \zeta'_{it}\delta\right)$ and then including the estimate of bias correction term in the second stage where we apply traditional SFA methods.[19] For the two-stage approach, the standard errors need to be corrected, e.g., by a bootstrap procedure. In our empirical section, we use the limited information maximum likelihood estimator that we presented in this section, i.e., the one-stage method.

Amsler et al. (2016, 2017) relax the conditional independence assumption for ε^s_{it} and $\tilde{u}_{it}$ by using a copula approach. Kutlu et al. (2018) show by simulations that if the firm-specific individual effects are included in the model, even if ε^s_{it} and $\tilde{u}_{it}$ are correlated conditionally, the estimates are still reasonable. When it is difficult to find instruments for endogenous variables, one may use the copula approach proposed by Tran and Tsionas (2015). Their model does not require the availability of outside information. Instead, to obtain the instruments, a flexible joint distribution of the endogenous variables and composed error is constructed. If the researcher is inclined to use Bayesian methods, the stochastic frontier model of Griffiths and Hajargasht (2016) can be applied to our framework. Traditional stochastic frontier models impose inefficient behavior on all firms. If it is believed there is a mixture of efficient and inefficient firms in the sample, it is possible to apply the model of Tran and Tsionas (2016).[20] Finally, one can disentangle firm specific heterogeneity from inefficiency by using variations of true-fixed effects (or true-random effects) model of Greene (2005a, 2005b) that allow endogeneity as in Kutlu et al. (2018). For example, in the airport and banking cost efficiency contexts Kutlu et al. (2018) and Kutlu and McCarthy (2016), respectively, illustrate that efficiency estimates can be substantially different if productive unit heterogeneity is not controlled in the estimations.

---

[18] We may replace $u_{it} = h_{it}\tilde{u}_{it}$ assumption by $u_{it} = h_{it}\tilde{u}_i$ so that $\tilde{u}_i$ is a firm specific term. This would be in line with the panel data stochastic frontier models.

[19] This bias correction term is said to be a control function.

[20] See also Guan et al. (2009) and Tran and Tsionas (2013) for GMM based approaches for handling endogeneity in SFA framework.

In the stochastic frontier model that we presented, we have exogenous and endogenous regressors along with some "outside instruments." Hence, our identification assumptions are somewhat different from the standard stochastic frontier models without endogenous variables. Our main identification assumption is that the exogenous variables (including the outside instruments) are uncorrelated with ε_{it}^s and w_{it} and that there are enough instruments. If ε_{it}^s and w_{it} are independent of the exogenous variables (including the outside instruments), then for each endogenous variable and its functions a single control function would be enough to achieve identification. For example, if z is a valid instrument for an endogenous variable x, then the model parameters can be identified by a single control function even when the model has x and x^2 as regressors.[21]

In order to make our contribution clearer, we finalize this section by comparing our model with two closely related papers. Kutlu and Sickles (2012) consider a dynamic conduct parameter model in which under the full market power scenario the firms play an efficient super-game equilibrium where the firms cooperate subject to incentive compatibility constraints. They estimate a market specific aggregate model and assume that the corresponding aggregate incentive compatibility constraint is a function of efficiency. The only place that efficiency enters their model is within the incentive compatibility constraint causing the parameter estimates from the static counterpart of their model to be invariant to the presence of inefficiency. This contrasts with our setting as our static model directly includes inefficiency in the supply equation. For their static model, the presence of efficiency matters only in the calculation of DWL. Although Kutlu and Sickles (2012) introduce the marginal cost efficiency concept, they assume that marginal cost efficiency equals cost efficiency (in the SFA sense), which is a somewhat strong assumption. Hence, they simply estimate a stochastic frontier cost function to obtain the cost efficiencies of the firms. A direct implication of this is that their dynamic model requires total cost data.

In another closely related study, Delis and Tsionas (2009) estimate a supply-demand-cost system where the cost function is modeled in the SFA framework. Hence, this study requires total cost data and efficiency concept is cost efficiency in the SFA sense. Their supply equation is derived from a standard conduct parameter model, which is invariant to the presence of inefficiency, and cost inefficiency enters their model only through the stochastic frontier cost function. This contrasts with our setting as our model directly includes (marginal cost) inefficiency in the supply equation. As a final remark, supply equation of Delis and Tsionas (2009) has revenue as the dependent variable

and the right-hand-side variables include cross products of (total) cost with many other variables, which are expected to be endogenous in this setting. This complication may pose some estimation related difficulties if not controlled properly.

4 The data

In order to testify our theoretical framework, we use the U.S. domestic airline data. One of the main data sources that we use is the Passenger Origin-Destination Survey of the U.S. Department of Transportation (DB1B data set). This data set is a 10% random sample of all tickets that originate in the U.S. on domestic flights. In our data set a market is defined as a directional city-pair (route). Calculation of prices and quantities are based on the tickets that have no more than three segments in each direction. Approximately 1% of tickets are eliminated during the elimination of tickets with more than 3 segments. We only focus on coach class tickets due to the differences in demand elasticities and other characteristics between coach class and high-end classes (first class and business class).

Our data set covers the time period from the first quarter of 1999 to fourth quarter of 2009. During this time period, the U.S. airlines faced serious financial problems. As pointed out by Duygun et al. (2016), the financial losses for domestic passenger airline operations during this time period were substantially higher than their losses between 1979 and 1999. Increase in taxes and jet fuel prices, relatively low fares, and sharp decrease in demand were some of the challenges facing U.S. airlines. During this time period, there were dramatic increases in load factors.

We provide the details about data construction process as follows. First, all multi-destination tickets are dropped as it is difficult to identify the ticket's origin and destination without knowing the exact purpose of the trip. Second, any itinerary that involves international flights was eliminated. Third, the fare class for high-end carrier was adjusted. That is, for some airlines, due to marketing strategies, only first class and business class (high-end) tickets are provided to consumers on all routes. However, the quality should be taken as coach class. For such airlines, if there is no coach class tickets from a certain carrier in a given quarter, we consider all tickets as coach class tickets. Fourth, tickets that have high-end segments and unknown fare classes were dropped.

Following Borenstein (1989) and Brueckner et al. (1992), we assume that ticketing carrier is the relevant airline. After further elimination of multi-ticketing-carrier tickets, firm specific average segment numbers (SEG) and average stage length (SL) on a given route are calculated as

---

[21] See Amsler et al. (2016) for further details of identification in SFA models with endogenous variables.

indicators for quality and costs. Moreover, our data set includes a distance variable which is the shortest directional flight distance (*DIST*). A ticket is online when the one-way ticket does not involve change of airplanes. The online variable is the percentage of online tickets.

For the price variable, we use the average price of all tickets for a given airline on a given route in given quarter.[22] All tickets with incredible[23] prices are dropped from our data set. Following Borenstein ([1989](#)) and Ito and Lee ([2007](#)), we eliminate the open-jaw tickets since it would be difficult to distribute the ticket price into outbound and inbound segment for open jaw tickets. We drop the tickets that have a price less than 25 dollars or higher than 99 percentile or more than 2.5 times standard deviation from the mean for each airline within a route. The tickets that have price less than 25 dollars are considered as frequent flyer program tickets and the tickets that have prices higher than 99 percentile are considered to be input (key punch) errors for the data set. For the round trip tickets, we divided the total price by two to get the one-way price.

The cost data set is constructed from the firm level data of DOT's airline production data set (based on Form 41 and *T*100).[24] While the airline-specific total cost data are available for the whole U.S. airline industry, the route-airline-specific total cost data are not available. We control for three types of important costs: labor price (*LP*), energy price (*EP*), and capital price (*KP*). The salaries and benefits for five main types of personnel are provided in Form 41/P6. Annual employee number is given in Form 41, P10. We interpolated the annual employee data to get the quarterly values. For energy price, we only capture the cost based on aircraft fuel. The energy input is developed by combining information on aircraft fuel gallons used with expense data per period. Flight capital is described by the average size (measured in number of seats) of the fleet. The number of aircraft that a carrier operated from each different model of aircraft in airline's fleet is collected from DOT Form 41. For each quarter, the average number of aircraft in service is calculated by dividing the total number of aircraft days for all aircraft types by the number of days in the quarter. This serves as an approximate measure of the size of fleet.

In order to estimate the demand, we also include the city specific demographic variables: per capita income (*PCI*) and population (*POP*). We obtained the city level per capita income and population data from Bureau of Labor Statistics. We interpolate the annual data to get the quarterly *PCI*

and population variables for each city. For each origin-destination city-pair, we use the population weighted *PCI* as the route-specific *PCI* measure. Similarly, city-pair population is the average population of origination and destination cities. In order to get the real prices, we deflated the nominal prices by Consumer Price Index (*CPI*) and use the first quarter of 1999 as the base time period. Because metropolitan areas have available demographic information whereas airports located in small cities do not, the number of the city-pairs is further reduced in our final data set.

We apply our theoretical method on the routes that originate from Chicago, which is a popular choice because of its relatively large airport and wide selection of airline firms. For instance, Brander and Zhang ([1990](#)) use 33 Chicago-based routes in their studies. In our final data set, we further eliminate the small firms and small routes. On a given route, small firms with market shares less than 0.01 are eliminated. For a given quarter, any route with enplanements less than 1800, i.e., 20 passengers per day or routes with less than 30 observations are dropped from the analysis. Table [1](#) presents the summary statistics for Chicago-based routes.

Low cost carrier (LCC) is a dummy variable that equals 1 if the ticketing carrier is a low cost carrier, otherwise it is 0. Number of firms represents the total number of firms that operate on the route. The total number of passengers is the total number of tickets sold on a given route by all airlines together in a given quarter. Total number of passengers for other routes (*OQOTH*) variable is the total number of tickets that are sold on the all other routes that share the same origination city. We use the geometric market share (*GEOS*) variable of Gerardi and Shapiro ([2009](#)) as an instrument.[25] Another instrumental variable that we use is *GEONF*. This variable is the product of *GEOS* and $nf^* = \sqrt{n_o * n_d}$, where n_o denotes to the mean value of number of firms for all routes that share the same origination as route of interest while the n_d refers to the mean value of number of firms for all routes that share the same destination city.

The final data set contains 108 routes that originate from Chicago and 14 carriers. The low cost carriers are Frontier Airlines, JetBlue Airways, Southwest Airlines, and Spirit Airlines. The remaining carriers are: Alaska Airlines, American Airlines, Continental Airlines, Delta Airlines, Northwest Airlines, United Airlines, US Airways, America West Airlines, ATA Airlines, and Trans World Airways.

---

[22] The average is calculated after eliminating the outliers.
[23] The incredible tickets are defined by DB1B data set.
[24] For more information about DOT's dataset see Weiher et al. ([2002](#)) and Good et al. ([2008](#)).

[25] *GEOS* is the *GENSP* variable that is used in Gerardi and Shapiro ([2009](#)). *GENSP* is similar to the *GEOSHARE* variable of Borenstein and Rose ([1994](#)). The difference is that Borenstein and Rose ([1994](#)) use average daily enplanements while we use average quarterly enplanements.

Table 1 Summary for Chicago Originating Routes

Variables: Name in estimation	Mean	S.D.	Min	Max
Low Cost Carrier Dummy: LCC	0.11225	0.31569	0	1
ln(Population): ln(POP)	15.50922	0.18392	15.33637	16.46417
ln(Per Capita Income): ln(PCI)	10.4226	0.03828	10.33827	10.61397
ln(Stage Length): ln(SL)	6.40837	0.59661	4.69135	8.34378
ln(Distance): ln(DIST)	6.63774	0.65719	4.69135	8.35303
ln(Average Fleet Size): ln(SIZE)	4.97262	0.08467	4.80562	5.33165
ln(Labor Price): ln(LP)	9.23198	0.7346	5.58453	10.00725
ln(Capital Price): ln(KP)	7.27984	0.76111	3.6658	8.1754
ln(Fuel Price): ln(FP)	3.87272	0.58681	2.37945	4.83215
Number of Firms: NF	8.2814	3.14474	1	24
ln(Price Per Ticket): ln(P)	4.96474	0.33975	3.72431	7.1672
ln(Number of Passengers): ln(Q)	9.79688	1.43261	7.49554	13.1454
ln(Number of Passengers for Other Routes): ln(OQOTH)	15.29761	0.10457	14.98133	15.47699
Market Share: s	0.22057	0.22458	0.01001	1
Geometric Market Share: GEOS	0.19418	0.12814	0.00682	1
Geometric Market Share*Number of Firms: GEONF	1.08933	0.66467	0.03784	4.27413
Online Rate: ONLINE	0.66744	0.36844	0	1
Top 4 Concentration Ratio: CR4	0.9195	0.11333	0.07516	1
ln(Average Number of Segments): ln(SEG)	0.3698	0.32171	0	1.09861
Number of Observations	18209			

5 Empirical example

The purpose of this section is providing an empirical example for our theoretical model. In particular, we estimate time-varying firm-route-specific conducts and marginal cost efficiencies of U.S. airlines for Chicago based routes. Like Brander and Zhang (1990) and Oum et al. (1993), we only consider coach class tickets. Brander and Zhang (1990) conclude that there is Cournot type competition in the airline industry, i.e., competition is quantity based. Hence, we assume a quantity based competition.[26] Our city-pair markets consist of one-way or round-trip directional trips having up to three segments in each direction. We divide the total ticket price by 2 to get the one-way fare for round-trip tickets. The demand and supply equations are estimated separately. The market demand equation is given by:

$$\ln P_{itr} = \beta_0 + \beta_1 \ln Q_{tr} + \beta_2 \ln PCI_{tr} \ln Q_{tr} + f_d(X_{d,itr}) + \varepsilon_{itr}^d \tag{11}$$

where f_d is a function of demand related variables, Q_{tr} is the total quantity at time t for route r, $X_{d,itr}$ is a vector of demand related variables, and ε_{itr}^d is the conventional two-sided error term. We assume that $\ln Q_{tr}$ and $\ln PCI_{tr} \ln Q_{tr}$

are endogenous. Along with the exogenous variables included in the demand model, our instrumental variables are $GEOS_{itr}$, $GEONF_{itr}$, $lnOQOTH_{rt}$, logarithm of labor price (ln LP_{it}), logarithm of capital price (ln KP_{it}), and logarithm of energy price (ln EP_{it}).

The supply equation is given by:

$$\ln P_{itr} = \ln c_{itr}^* + g_{itr} + u_{itr} + \varepsilon_{itr}^s \tag{12}$$

where c_{itr}^* is the marginal cost when firms achieve full efficiency, $g_{itr} = -\ln\left(1 - \frac{s_{itr}}{E_{tr}}\theta_{itr}\right)$ is the market power term, $u_{itr} \geq 0$ is the inefficiency term, and ε_{itr}^s is the conventional two sided error term. The parameters of the E_{tr} term is identified through the demand equation. We assume that the efficient marginal cost, c_{itr}^*, is constant with respect to quantity, i.e., it is not a function of quantity but maybe a function of exogenous cost shifters. Hence, as we described in the theoretical model section, the theoretical values for cost and marginal cost efficiencies coincide in this model. Although constant marginal cost is a relatively strong assumption, it is commonly used in the conduct parameter models. Iwata (1974), Genesove and Mullin (1998), Corts (1999), and Puller (2007) exemplify some papers that use this assumption in a variety of conduct parameter settings. We use this simplifying assumption to illustrate our methodology. Nevertheless, we approximate the efficient marginal cost function by a fairly flexible function of input prices and

---

[26] Kutlu and Sickles (2012) also assume quantity competition for airlines.

other cost related exogenous variables. These cost related variables include year, quarter, and airline dummy variables, which capture time-firm-specific unobserved factors.[27] Moreover, constant marginal cost assumption is not unreasonable at least around the equilibrium as there is substantial empirical evidence supporting constant returns to scale for the airline industry. We model g_{itr} as in the theoretical model section and assume that $X_{g,itr} = (s_{itr}, CR_{4,tr}, \ln DIST_r, t, E_{tr}, 1)'$ where $CR_{4,tr}$ is the concentration ratio for largest four firms on route r at time t. We assume that $u_{itr} = h_{itr}\tilde{u}_{itr}$ and $\tilde{u}_{itr} \sim \mathbf{N}^+(0, \sigma_u^2)$ where $\sigma_u^2 = \exp\left(X'_{g,itr}\beta_u\right)$; and $\varepsilon_{itr}^s \sim \mathbf{N}(0, \sigma_\varepsilon^2)$ where $\sigma_\varepsilon^2 = \exp(\beta_\varepsilon)$. For the supply side, s_{itr} and $CR_{4,tr}$ are assumed to be endogenous. Our instrumental variables are $GEOS_{itr}$, $GEONF_{itr}$, $\ln POP_{tr}$, and $\ln PCI_{tr}$. The estimations of the supply relations are done by using the limited information maximum likelihood estimation method that we described in our theoretical model section.

6 Results

In this section, we present our estimation results. The demand parameter estimates for the routes originating from Chicago are given in Table 2. We estimated the inverse demand equation by 2SLS. Our demand model controls for year, quarter, and airline. The demand elasticities are negative at each observation, i.e., $E_{tr} > 0$.[28]

For the supply function, as we described previously, we use the one-stage limited information maximum likelihood approach to deal with endogeneity. In order to illustrate the consequences of ignoring marginal cost inefficiency, we estimated two supply models: the first one allows inefficiency (benchmark model) and the second one assumes full efficiency so that $u_{itr} = 0$ (full-efficiency model). The full-efficiency model is a standard conduct parameter model, which helps us to compare our benchmark estimates with the standard conduct parameter models. Both models include airline, year, and quarter dummy variables.[29]

Table 3 shows the estimation results. The bias correction terms (η) are jointly significant at any conventional significance level, which is an indication of endogeneity. The median of the conduct estimates from the benchmark model

Table 2 Estimation for demand function

Dependent variable: price	Estimates	Std. err.
ln(Q)	2.91356***	(0.50629)
ln(Q) ln(PCI)	−0.29270***	(0.04861)
ONLINE	0.01940*	(0.00880)
ln(DIST)	7.47912***	(0.95064)
ln(SEG)	−10.03281***	(0.97631)
ln(SIZE)	−15.81152***	(2.31769)
ln(SL)	−8.24091***	(0.97745)
ln(DIST) Square	−0.22445**	(0.07828)
ln(SEG) Square	0.53698***	(0.11695)
ln(SIZE) Square	1.50167***	(0.22554)
ln(SL) Square	−0.62278***	(0.08377)
ln(DIST) ln(SEG)	−0.65458**	(0.19646)
ln(DIST) ln(SIZE)	−1.93435***	(0.18417)
ln(DIST) ln(SL)	0.90929***	(0.15754)
ln(SEG) ln(SIZE)	2.44755***	(0.19223)
ln(SEG) ln(SL)	0.24996	(0.19551)
ln(SIZE) ln(SL)	1.98114***	(0.18886)
ln(POP)	0.30783***	(0.01673)
ln(PCI)	4.19634***	(0.56805)
LCC	−0.21325***	(0.00862)
Quarter Dummies	Yes	
Year Dummies	Yes	
Firm Dummies	Yes	
Centered R Square	0.4976	
Number of Observations	18209	

+ p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001

Robust standard errors are given in bracket

is 0.63, which is lower than the theoretical conduct value for Cournot competition, 1.[30] At the median, the extent of competition lies somewhere between perfect competition and Cournot competition. The median of conduct estimates from the full-efficiency model is somewhat lower, 0.22, suggesting a competitive market. In other words, when we allow inefficiency, the median conduct is closer to Cournot competition benchmark, i.e., $\theta = 1$; and when we assume full efficiency, the median conduct is closer to perfect competition benchmark, i.e., $\theta = 0$. Moreover, the Kolmogorov-Smirmov test rejects the equality of the distributions at any conventional significance levels. This illustrates the importance of considering marginal cost inefficiency when estimating a conduct parameter model.

LCCs, due to their special operating style,[31] tend to have lower marginal costs compared to other airlines, which helps

---

[27] If the airlines are playing a version of dynamic conduct parameters game that is suggested by Puller (2009), route-specific time dummy variables would capture dynamic factors that enter the airlines' optimization problems as well. In this case, the estimates of parameters, conduct parameters, and efficiencies would still be consistent. However, the marginal cost estimates may be downward biased as the prediction of marginal costs include these dynamic factors. Nevertheless, the efficient full marginal cost estimates would be consistent.

[28] Recall that we define $E_t = -\frac{\partial Q_t}{\partial P_t}\frac{P_t}{Q_t}$.

[29] Hence, as we illustrate in the Appendix, our parameter and efficiency estimates are consistent even when the marginal costs are stochastic.

[30] The median of theoretical conduct values for joint profit maximization scenario is 6.77, which is the median of $\frac{1}{s_{itr}}$.

[31] For example, some of them operate only on certain routes in order to reduce costs.

Table 3 Estimation for supply function

Supply function: price	Inefficiency allowed		Full efficiency	
	Estimates	Std. err.	Estimates	Std. err.
ln(KP)	−0.72825***	(0.13523)	−0.63778***	(0.15070)
ln(FP)	0.02501	(0.11625)	0.07255	(0.12892)
ln(LP)	−0.34062*	(0.14423)	−0.23348	(0.16666)
ln(KP) Square	0.02449	(0.02310)	0.01674	(0.02643)
ln(FP) Square	0.01010	(0.01061)	0.01876	(0.01164)
ln(LP) Square	0.02082	(0.01929)	0.01612	(0.02175)
ln(KP) ln(FP)	0.02793+	(0.01550)	0.02603	(0.01717)
ln(KP) ln(LP)	0.02922	(0.04268)	0.03147	(0.04818)
ln(FP) ln(LP)	−0.04628**	(0.01768)	−0.05404**	(0.01982)
ONLINE	−0.11235***	(0.00769)	−0.13278***	(0.00853)
ln(DIST)	3.25569***	(0.96877)	1.68620	(1.05468)
ln(SEG)	−3.81246***	(0.95276)	−1.65408	(1.01118)
ln(SIZE)	−15.47651***	(3.14511)	−6.51047+	(3.37444)
ln(SL)	−3.46404***	(1.01353)	−1.90943+	(1.10320)
ln(DIST) Square	0.41059***	(0.07352)	0.35812***	(0.07470)
ln(SEG) Square	0.88206***	(0.11481)	0.81164***	(0.11475)
ln(SIZE) Square	1.63611***	(0.31323)	0.75112*	(0.33609)
ln(SL) Square	0.19687*	(0.08723)	0.11232	(0.09075)
ln(DIST) ln(SEG)	−2.11495***	(0.18315)	−1.75739***	(0.17727)
ln(DIST) ln(SIZE)	−0.76913***	(0.19136)	−0.55407**	(0.21109)
ln(DIST) ln(SL)	−0.51857**	(0.15756)	−0.37967*	(0.16206)
ln(SEG) ln(SL)	1.93489***	(0.18700)	1.50137***	(0.18227)
ln(SEG) ln(SIZE)	0.94664***	(0.18911)	0.60553**	(0.20198)
ln(SIZE) ln(SL)	0.64006**	(0.20052)	0.40393+	(0.22051)
LCC	−0.76658***	(0.07363)	−0.76248***	(0.08575)
Quarter Dummies	Yes		Yes	
Year Dummies	Yes		Yes	
Firm Dummies	Yes		Yes	

Nonlinear function: g	Inefficiency allowed		Full efficiency	
	Estimates	Std. err.	Estimates	Std. err.
s	17.62885***	(1.77711)	17.86643***	(1.87244)
CR4	8.82176***	(2.35791)	16.89555***	(3.57813)
ln(DIST)	0.38011	(0.33349)	0.47333+	(0.27124)
t	−0.22842***	(0.02998)	−0.23597***	(0.03529)
Elasticity of Demand	−0.13075	(0.28139)	−0.89407**	(0.29823)
Constant	−10.53598*	(5.06577)	−13.87330**	(4.77015)
σ_w				
Constant	−3.46254***	(0.03206)	−2.70253***	(0.01051)
σ_u				
s	0.41774***	(0.08396)		
CR4	3.95924***	(0.28411)		
ln(DIST)	−0.94727***	(0.04206)		
t	−0.03700***	(0.00233)		
Elasticity of Demand	−0.15413***	(0.04344)		
Constant	1.94036***	(0.53855)		

Table 3 (continued)

Nonlinear function: g	Inefficiency allowed		Full efficiency	
	Estimates	Std. err.	Estimates	Std. err.
η (bias correction term)				
s	−0.33883***	(0.02233)	−0.33863***	(0.01947)
CR4	0.09003**	(0.02838)	0.58208***	(0.02438)
Log-likelihood	30772.7323		30033.13415	
Number of Observations	18209		18209	

+p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001

them to offer lower fares. Hence, it is worthwhile to examine the decomposition of conducts based on LCCs and non-LCCs. Our conduct estimates from the benchmark model for LCC and non-LCC carriers are 0.24 and 0.74, respectively. Therefore, while the conducts of LCCs are closer to perfectly competitive values, the conducts of non-LCCs are closer to Cournot competition values. The corresponding estimates from the full-efficiency model are 0.11 and 0.25, indicating relatively competitive markets for both LCCs and non-LCCs. Hence, ignoring inefficiency leads to underestimation of conducts for both LCCs and non-LCCs. The underestimation of welfare loss is boosted by the fact that the full-efficiency model ignores the marginal cost inefficiency.

The conduct parameter estimates show that an airline with higher market share tends to have higher market power. In markets with high CR_4 values, it may be easier for airlines with higher market share to cooperate. The positive coefficient of CR_4 in conduct verifies this. For the time period that we examine, the U.S. airlines seem to be losing market power over time. For longer flight distances the alternative transportation means (e.g., bus or car) are likely to become less attractive to the consumers. This reduction in outside competition suggests a positive relationship between market power and distance. The positive coefficient of distance variable for the market power term is in line with this intuition.

In Figs. 2 and 3, the efficiency and conduct estimates from our benchmark model are presented. In our benchmark estimates, the median efficiency estimates for the whole sample and non-LCC carriers are 82.6 and 84.4%, respectively. Hence, the efficiencies of LCCs and non-LCCs are similar. The parameter estimates for the inefficiency term show that an airline with higher market share tends to have higher inefficiency. Moreover, higher CR_4 values lead to lower efficiency and the correlation between conduct parameter and efficiency is −0.19, which is statistically significant at any conventional significance level. These are in line with the QLH which postulates that higher market
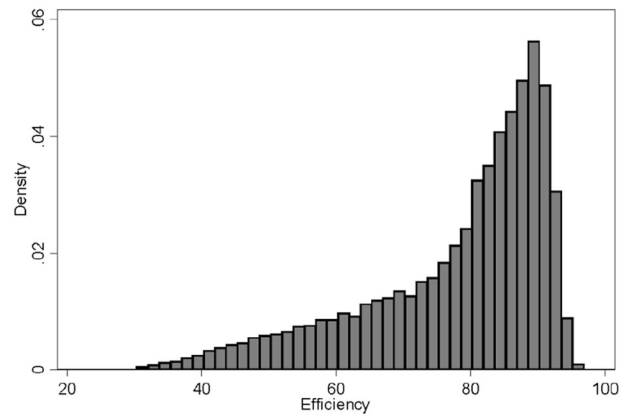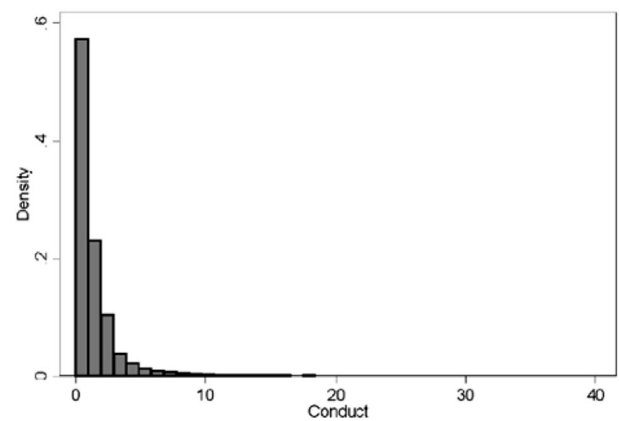

Fig. 2 Efficiency estimates

Fig. 3 Conduct estimates

power leads to lower efficiency levels.[32] Based on our benchmark model, the medians of price-marginal cost markups, price-efficient marginal cost markups, and prices are $4.65, $30.32, and $140.46, respectively. However, for the full-efficiency model, the median of price-marginal cost markups is $2.49. Historically, airlines have been

---

[32] Note that since u has a half normal distribution its mean depends on σ_u^2. In particular, the mean of u is an increasing function of σ_u^2.

challenged in their efforts to generate high profits. These markup values from our benchmark model indicate that airlines may partially be responsible for the financial difficulties that they face. Our study shows that the answer to achieving reasonable profit levels may be through improving efficiency.

Finally, we calculate a lower bound for bias in DWL estimates when the researchers use the full-efficiency model instead of the benchmark model. This number is obtained by estimating the shaded rectangular area between MC, EMC, and Q_θ (see Fig. 1) where we assume that Q_θ equals total observed quantity on a given route in a given quarter. The median of the lower bound over all routes and quarters is \$427,715. This area measures the DWL that is caused by efficiency loss only. Hence, the DWL due to misallocation of quantity that stems from efficiency loss is not included.[33] Nevertheless, this lower bound clearly illustrates the severe consequences of ignoring marginal cost inefficiency when evaluating welfare loss using the conventional conduct parameter models.

7 Summary and concluding remarks

In this paper, we provided a conduct parameter based framework to estimate market powers and (marginal cost) efficiencies of firms simultaneously. Our methodology enables us to relax the total cost data requirement for stochastic frontier models. Total cost data may not be available for a variety of reasons. For example, firms might not want to reveal this potentially strategic information. Even when some form of total cost data are available, the data may not reflect the total cost of the relevant unit that we want to examine. For instance, in our empirical example, for the U.S. airlines only firm specific total cost data is available for the entire U.S. airline system. In such cases, the conventional stochastic frontier models cannot estimate firm-route specific efficiencies as this would require firm-route specific total cost data.

Besides relaxing a vital data requirement, our methodology aims to overcome some estimation issues. Efficiencies are generally measured by the distance between the units of production and the best practice units observed in the market. If the performance of the best-practice units depends on their market powers, then the efficiency estimates that are not taking this into account would not be accurate. We overcome this difficulty by explicitly modeling a conduct parameter game in an environment where firms are allowed to be inefficient.

In the Appendix, we provide extensions of our model for a variety of different settings including capacity constraints, stochastic marginal costs, multi-output firms, and dynamic strategic interactions. Another potential extension of our conduct parameter model is so that the firms price discriminate. For example, the marginal cost efficiency concept can be incorporated into the conduct parameter games of Kutlu (2012a, 2017a) and Kutlu and Sickles (2017). Such an extension would enable us to understand the connection between price discrimination, market power, and efficiency better.[34] Hence, our theoretical model serves as a guideline as to how conduct parameter and efficiency can be estimated simultaneously without requiring total cost data. Our guideline can be applied to a variety of conduct parameter settings.

As for the standard DWL calculations, our efficiency measure for marginal cost does not consider the fixed costs for the short run. It is possible to consider dynamic frameworks where prior investments and R&D may affect fixed costs and marginal costs. Hence, suboptimal investment decisions may result in suboptimal marginal cost levels even when the marginal cost may seem optimal in a single time horizon for given investment and R&D levels. This type of efficiency is not controlled in our conduct parameter model or in a standard stochastic frontier model. Moreover, similar to the standard stochastic frontier models, we assume that the input market is perfectly competitive. Hence, our model ignores potential market power in the input markets. The advantage of our framework over SFA setting is that we can make such structural extensions relatively easier compared to SFA setting using the already-established industrial organization literature.

We applied our methodology to estimate the conducts and marginal cost efficiencies of the U.S. airlines for the Chicago based routes between 1999I-2009IV. We found that the market concentration and market share of airlines are negatively related to the marginal cost efficiency, which is in line with the QLH. We also found that the conduct and DWL estimations may be seriously distorted if inefficiency is ignored. A more extensive empirical study is warranted but we consider this outside the scope of this paper. For example, a future study may consider a more complete list of U.S. routes and a longer time period. Moreover, although we control for the airline specific factors by including airline and time dummy variables, a study including variables related to on time rate, food service, aircraft engine information, and aircraft seat specification may be worth exploring.

---

[33] Note that the triangular welfare loss area (shaded area between Q_θ and Q_C in Fig. 1) for the benchmark model is larger compared to that of full-efficiency model. Therefore, the rectangular area that we estimated is a lower bound for the DWL bias.

[34] Borenstein and Rose (1994), Stavins (2001), Gerardi and Shapiro (2009), and Chakrabarty and Kutlu (2014) examine the relationship between market power and price discrimination.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

8 Appendix: extensions of theoretical model

The theoretical model we presented illustrates how the marginal cost efficiency concept can be incorporated to a simple yet commonly used conduct parameter model. It is possible to apply similar ideas to a variety of different conduct parameter models. In this appendix, we briefly present some alternative frameworks where marginal cost efficiency concept can be used.

8.1 Capacity constraints

Our model can be extended to a setting in which firms have capacity constraints. This extension of our model is inspired by the conduct parameter model proposed by Puller (2007). In the presence of capacity constraints the optimization problem for firm i becomes:

$$\max_{q_{it}} P_t q_{it} - C_{it} \, \text{s.t.} \, q_{it} \leq K_{it} \tag{13}$$

where K_{it} is the capacity constraint that firm i is facing at time t. The first order conditions for the corresponding conduct parameter game are:

$$P_t = -\theta_{it} P_t' q_{it} + c_{it} + \lambda_{it} \tag{14}$$

where $\lambda_{it} \geq 0$ is the shadow cost of capacity which can be estimated by including variables capturing extent of capacity constraints. For example, Puller (2007) uses a dummy variable, which equals one when a constraint is binding. If we let $\tilde{c}_{it}^* = c_{it}^* + \lambda_{it}$ and $\tilde{c}_{it} = c_{it} + \lambda_{it}$, as earlier we have:

$$\ln P_{it} = \ln \tilde{c}_{it}^* + g_{it} + u_{it} + \varepsilon_{it}^s. \tag{15}$$

The presence of λ_{it} may make it difficult to estimate this model. A solution may be to use an inefficiency variable, $r_{it} \geq 0$, which enters the equation additively:

$$P_{it} = -\theta_{it} P_t' q_{it} + c_{it}^* + \lambda_{it} + r_{it} + \varepsilon_{it}^s \tag{16}$$

where $c_{it} = c_{it}^* + r_{it}$, $r_{it} \sim N^+(\mu_r, \sigma_r^2)$, and ε_{it}^s is the usual two-sided error term.

In line with the standard stochastic frontier models, in our original model the realized marginal cost (c_{it}) is assumed to be a multiple of minimum (frontier) marginal cost $\left(c_{it}^*\right)$; and the ratio of efficient marginal cost to realized marginal cost $\left(c_{it}^*/c_{it}\right)$ is defined as marginal cost efficiency. In this setting, the realized marginal cost is assumed to follow $\ln c_{it} = \ln c_{it}^* + u_{it}$ where $u_{it} \geq 0$.[35] Efficiency is calculated using the estimates from this log-transformed multiplicative form by $Eff_{it} = c_{it}^*/c_{it} = \exp(-u_{it})$. In this example, the inefficiency term, r_{it}, enters the model additively. Hence, r_{it} measures the marginal cost efficiency additively. That is, we model the realized marginal cost by $c_{it} = c_{it}^* + r_{it}$ where $r_{it} \geq 0$. Similar to its multiplicative counterpart, this model additively measures how large the realized marginal cost is relative to the efficient marginal cost. After estimating the parameters of the additive model, efficiency can be estimated by $Eff_{it} = c_{it}^*/c_{it} = c_{it}^*/\left(c_{it}^* + r_{it}\right)$. Depending on the structural model used, the researcher may find either multiplicative, u_{it}, or additive, r_{it}, versions of efficiency variable convenient. However, it seems that for the game theoretic models, the additive version may be more frequently needed.

8.2 Stochastic marginal cost

Until now we considered models where the firms have perfect information about the marginal costs. However, firms may not always have perfect information about their marginal costs. In such cases, they would maximize their expected profits. We provide a simple example for how this issue can be handled in our framework. Assume that the marginal costs of the firms are stochastic in the sense that:

$$\ln c_{it} = \ln c_{it}^* + u_{it} + v_{it} \tag{17}$$

where the inefficiency level, u_{it}, and the distribution of $v_{it} \sim N\left(0, \sigma_v^2\right)$ is known by the firms but v_{it} is not observed by neither the firms nor the researcher. In this scenario, the perceived marginal revenue would be equal to the expected marginal cost. As earlier, the perceived marginal revenue is given by:

$$PMR(\theta_{it}) = P_t \left(1 - \frac{s_{it}}{E_t} \theta_{it}\right). \tag{18}$$

The expectation of marginal cost is given by:

$$\begin{aligned} E[c_{it}] &= c_{it}^* \exp(u_{it}) E[\exp(v_{it})] \\ &= c_{it}^* \exp(u_{it}) \exp\left(\tfrac{1}{2}\sigma_v^2\right) \Rightarrow \\ \ln(E[c_{it}]) &= \gamma + \ln c_{it}^* + u_{it} \end{aligned} \tag{19}$$

---

[35] For expositional purposes, we ignore the two-sided error term.

where $\gamma = \frac{1}{2}\sigma_v^2$. Hence, after adding the error term, the supply equation for firm i is given by:

$$\ln P_{it} = \gamma + \ln c_{it}^* + g_{it} + u_{it} + \varepsilon_{it}^s. \tag{20}$$

This supply equation is the same as the deterministic cost function scenario except for the addition of the constant term, γ. If v_{it} is heteroskedastic so that $v_{it} \sim N(0, \sigma_{vi}^2)$ where σ_{vi}^2 is firm specific, the model becomes:

$$\ln P_{it} = \gamma_i + \ln c_{it}^* + g_{it} + u_{it} + \varepsilon_{it}^s. \tag{21}$$

This model can be estimated using the true individual effects model of Kutlu et al. (2018).

8.3 Multi-output firms

Single-output scenario may be restrictive in some contexts such as banking (e.g., Berger and Hannan 1998; Koetter et al. 2012; Kutlu 2012b). Therefore, we provide a conduct parameter model with multi-output firms. Without loss of generality, we assume that there are two outputs and the corresponding demands are represented by $P_1(Q)$ and $P_2(Q)$ where $Q = (Q_1, Q_2)$ is the market output vector. The cost function for firm i is $C_i(q_i)$ where $q_i = (q_{1i}, q_{2i})$ represents the firm output vector. The profit function for firm i is given by:

$$\pi_{it} = P_1(Q_t)q_{1it} + P_2(Q_t)q_{2it} - C_{it}(q_i). \tag{22}$$

Hence, perceived marginal revenues for outputs are:

$$\begin{aligned} PMR_1(\theta_{it}) &= P_{1t} + \frac{\partial P_{1t}}{\partial Q_{1t}}\frac{\partial Q_{1t}}{\partial q_{1it}}q_{1it} + \frac{\partial P_{2t}}{\partial Q_{1t}}\frac{\partial Q_{1t}}{\partial q_{1it}}q_{2it} \\ &= P_{1t} + \left(\frac{\partial P_{1t}}{\partial Q_{1t}}q_{1it} + \frac{\partial P_{2t}}{\partial Q_{1t}}q_{2it}\right)\theta_{1it} \\ PMR_2(\theta_{it}) &= P_{2t} + \left(\frac{\partial P_{1t}}{\partial Q_{2t}}q_{1it} + \frac{\partial P_{2t}}{\partial Q_{2t}}q_{2it}\right)\theta_{2it} \end{aligned} \tag{23}$$

where $\theta_{it} = (\theta_{1it}, \theta_{2it}) = \left(\frac{\partial Q_{1t}}{\partial q_{1it}}, \frac{\partial Q_{2t}}{\partial q_{2it}}\right)$ is the vector of conducts for each output. In line with Nash equilibrium solution, we assume that $\frac{\partial Q_{2t}}{\partial q_{1it}} = 0$ and $\frac{\partial Q_{1t}}{\partial q_{2it}} = 0$. After adding the error terms, the supply relations become:

$$\begin{aligned} P_{1t} &= -\left(\frac{\partial P_{1t}}{\partial Q_{1t}}q_{1it} + \frac{\partial P_{2t}}{\partial Q_{1t}}q_{2it}\right)\theta_{1it} + c_{1it}^* + r_{1it} + \varepsilon_{1it}^s \\ P_{2t} &= -\left(\frac{\partial P_{1t}}{\partial Q_{2t}}q_{1it} + \frac{\partial P_{2t}}{\partial Q_{2t}}q_{2it}\right)\theta_{2it} + c_{2it}^* + r_{2it} + \varepsilon_{2it}^s \end{aligned} \tag{24}$$

where $r_{1it} \sim N^+\left(\mu_{r_1}, \sigma_{r_1}^2\right)$ and $r_{2it} \sim N^+\left(\mu_{r_2}, \sigma_{r_2}^2\right)$ represent the corresponding marginal cost inefficiencies additively so that $c_{1it} = c_{1it}^* + r_{1it}$ and $c_{2it} = c_{2it}^* + r_{2it}$; and ε_{1it}^s and ε_{2it}^s are the usual two-sided error terms. After estimating the demand equations, these supply equations can be estimated separately using stochastic frontier methods that we described above.

8.4 Dynamic strategic interactions

A formal treatment of conduct parameter games in which the strategic interactions of the firms are dynamic is beyond the scope of this study. However, following Puller (2009), we recommend including time fixed-effects, which may condition out the dynamic effects in firms' optimization problems.[36] Even though the estimates of parameters (including parameters of the conduct and efficiency) are consistent in this dynamic game scenario, we cannot separately identify the efficient marginal costs, c_{it}^*, and dynamic correction terms because the time dummies capture not only cost related unobserved factors that change over time but also the dynamic correction terms.[37] Nevertheless, except for the portion of time fixed-effects that contribute to c_{it}^*, the other parameters of c_{it}^* are identified. Moreover, many times c_{it}^* is not the main interest. In our empirical example, we assume a static model, which is not subject to these identification issues.

Finally, it is possible to extend the dynamic model of Kutlu and Sickles (2012) using similar procedures that we presented. However, since their game theoretical model estimates market-time specific conducts, the extension of this model would estimate market-time specific conduct and marginal cost efficiency.

References

Aigner DJ, Lovell CAK, Schmidt PJ (1977) Formulation and estimation of stochastic frontier production function models. J Econom 6:21–37

Amsler C, Prokhorov A, Schmidt P (2016) Endogeneity in stochastic frontier models. J Econom 190:280–288

Amsler C, Prokhorov A, Schmidt P (2017) Endogenous environmental variables in stochastic frontier models. J Econom 199:131–140

Berger A, Hannan TH (1998) The efficiency cost of market power in the banking industry: a test of the "quiet life" and related hypothesis. Rev Econ Stat 80:454–465

Berry S, Jia P (2010) Tracing the woes: an empirical analysis of the airline industry. Am Econ J Microecon 2:1–43

Borenstein S (1989) Hubs and high fares: dominance and market power in the U.S. airline industry. Rand J Econ 20:344–365

Borenstein S, Rose NL (1994) Competition and price dispersion in the U.S. airline industry. J Polit Econ 102:653–683

---

[36] See Puller (2009) for further details about his model and restrictions. One particular assumption that Puller (2009) makes is that the firms play an efficient super-game equilibrium when they cooperate. That is, they maximize the joint profit subject to incentive compatibility constraints. Hence, the corresponding efficient super-game equilibrium values are benchmark for the full market power case.

[37] In the static setting we don't have this identification issue as the dynamic correction terms are zero.

Brander JA, Zhang A (1990) Market conduct in the airline industry: an empirical study. Rand J Econ 21:567–583

Bresnahan TF (1982) The oligopoly solution is identified. Econ Lett 10:87–92

Bresnahan TF (1989) Studies of industries with market power, the handbook of industrial organization. North-Holland, Amsterdam

Brueckner J, Dyer N, Spiller PT (1992) Fare determination in airline hub-and-spoke networks. Rand J Econ 23:309–333

Chakrabarty D, Kutlu L (2014) Competition and price discrimination in the airline market. Appl Econ 46:3421–3436

Comanor WS, Leibenstein H (1969) Allocative efficiency, X-efficiency and the measurement of welfare losses. Economica 36:304–309

Cornwell C, Schmidt P, Sickles RC (1990) Production frontiers with time-series variation in efficiency levels. J Econom 46:185–200

Corts KS (1999) Conduct parameters and the measurement of market power. J Econom 88:227–250

Delis MD, Tsionas EG (2009) The joint estimation of bank-level market power and efficiency. J Bank Financ 33:1842–1850

Demsetz H (1973) Industry structure, market rivalry, and public policy. J Law Econ 16:1–9

Duygun M, Kutlu L, Sickles RC (2016) Measuring productivity and efficiency: a Kalman filter approach. J Product Anal 46:155–167

Genesove D, Mullin W (1998) Testing static oligopoly models: conduct and cost in the sugar industry, 1890–1914. Rand J Econ 29:355–377

Gerardi KS, Shapiro AH (2009) Does competition reduce price dispersion? New evidence from the airline industry. J Political Econ 107:1–37

Good D, Sickles RC, Weiher J (2008) A hedonic price index for airline travel. Rev Income Wealth 54:438–465

Greene WH (1980a) Maximum likelihood estimation of econometric frontier functions. J Econom 13:27–56

Greene WH (1980b) On the estimation of a flexible frontier production model. J Econom 3:101–115

Greene WH (2003) Simulated likelihood estimation of the normal-gamma stochastic frontier function. J Product Anal 19:179–190

Greene WH (2005a) Fixed and random effects in stochastic frontier models. J Product Anal 23:7–32

Greene WH (2005b) Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. J Econom 126:269–303

Griffiths WE, Hajargasht G (2016) Some models for stochastic frontiers with endogeneity. J Econom 190:341–348

Guan Z, Kumbhakar SC, Myers RJ, Lansink AO (2009) Measuring excess capital capacity in agricultural production. Am J Agric Econ 91:765–776

Hicks JR (1935) Annual survey of economic theory: the theory of monopoly. Econometrica 3:1–20

Ito H, Lee D (2007) Domestic code sharing, alliances, and airfares in the U.S. airline industry. J Law Econ 50:355–380

Iwata G (1974) Measurement of conjectural variations in oligopoly. Econometrica 42:947–966

Karakaplan MU, Kutlu L (2017a) Handling endogeneity in stochastic frontier analysis. Econ Bull 37:889–901

Karakaplan MU, Kutlu L (2017b) Endogeneity in panel stochastic frontier models: an application to the japanese cotton spinning industry. Appl Econ 49:5935–5939

Koetter M, Kolari JW, Spierdijk L (2012) Enjoying the quiet life under deregulation? Evidence from adjusted Lerner Indices for US Banks. Rev Econ Stat 94:462–480

Koetter M, Poghosyan T (2009) The identification of technology regimes in banking: implications for the market power-fragility nexus. J Bank Financ 33:1413–1422

Kumbhakar SC, Lovell CAK (2003) Stochastic frontier analysis. Cambridge University Press, Cambridge

Kutlu L (2010) Battese-Coelli estimator with endogenous regressors. Econ Lett 109:79–81

Kutlu L (2012a) Price discrimination in Cournot Competition. Econ Lett 117:540–543

Kutlu L (2012b) U.S. banking efficiency, 1984–1995. Econ Lett 117:53–56

Kutlu L (2017a) A conduct parameter model of price discrimination. Scott J Polit Econ 64:530–536

Kutlu L (2017b) A constrained state space approach for estimating firm efficiency. Econ Lett 152:54–56

Kutlu L, McCarthy P (2016) US airport governance and efficiency. Transp Res Part E 89:117–132

Kutlu L, Sickles CR (2012) Estimation of market power in the presence of firm level inefficiencies. J Econom 168:141–155

Kutlu L, Sickles CR (2017) Measuring market power when firms price discriminate. Empir Econ 53:287–305

Kutlu L, Tran K, Tsionas EG (2018) A time-varying true individual effects model with endogenous regressors, Working paper. https://ssrn.com/abstract=2721864

Lau LJ (1982) On identifying the degree of competitiveness from industry price and output data. Econ Lett 10:93–99

Lee C-Y, Johnson AL (2015) Measuring efficiency in imperfectly competitive markets: an example of rational inefficiency. J Optim Theory Appl 164:702–722

Lerner AP (1934) The concept of monopoly and the measurement of monopoly power. Rev Econ Stud 1:157–175

Maudos J, Fernandez de Guevara J (2007) The cost of market power in banking: social welfare loss vs. cost efficiency. J Bank Financ 31:2103–2125

Meeusen W, van den Broeck J (1977) Efficiency estimation from Cobb-Douglas production function with composed error. Int Econ Rev 8:435–444

Oum TH, Zhang A, Zhang Y (1993) Inter-firm rivalry and firm-specific price elasticities in deregulated airline markets. J Transp Econ Policy 27:191–192

Perloff JM, Karp LS, Golan A (2007) Estimating market power and strategies. Cambridge University Press, Cambridge

Perloff JM, Shen EZ (2012) Collinearity in linear structural models of market power. Rev Ind Organ 40:131–138

Puller SL (2007) Pricing and firm conduct in California's deregulated electricity market. Rev Econ Stat 89:75–87

Puller SL (2009) Estimation of competitive conduct when firms are efficiently colluding: addressing the Corts critique. Appl Econ Lett 16:1497–1500

Sickles RC (2005) Panel estimators and the identification of firm-specific efficiency levels in semi-pcolluding: addressing the Corts critiquearametric and non-parametric settings. J Econom 126:305–324

Stavins J (2001) Price discrimination in the airline market: the effect of market concentration. Rev Econ Stat 83:200–202

Stevenson RE (1980) Likelihood functions for generalized stochastic frontier estimation. J Econom 13:57–66

Tran KC, Tsionas EG (2013) GMM estimation of stochastic frontier model with endogenous regressors. Econ Lett 118:233–236

Tran KC, Tsionas EG (2015) Endogeneity in stochastic frontier models: copula approach without external instruments. Econ Lett 133:85–88

Tran KC, Tsionas EG (2016) On the estimation of zero-inefficiency stochastic frontier models with endogenous regressors. Econ Lett 147:19–22

Weiher JC, Sickles RC, Perloff JM (2002) Market power in the US airline industry. In: Slottje DJ (ed) Measuring Market Power. Emerald Group Publishing, North-Holland