TECHNICAL APPENDIX
"PRICE DISCRIMINATION AND MARKET POWER"

Not Intended for Publication

Shane Carbonneau, R. Preston McAfee, Hugo Mialon, and Sue Mialon

We are interested in statistically testing the difference between the coefficients of variation for the hub-to-spoke and spoke-to-hub markets of every airport pair in our sample. In this appendix, we develop two tests, one using the Delta Method, and one using the bootstrapping method, which are described in turn, after some preliminary notation is introduced.

Notation

Let μ denote population mean fare and let σ^2 denote the population variance for the fare. Then the coefficient of variation (CV) can be calculated simply as:

$$\theta = \frac{\sigma}{\mu}$$

Since population figures for μ and σ are not known, we make estimates using the observed sample of n observations. Given the univariate random sample of the fare $X = (x_1 \ , \ x_2 \ ,\ldots , x_n)$ drawn from the population, we can calculate several descriptive statistics. The sample estimate of the mean is calculated $\overline{x} = \frac{\sum_i x_i}{n}$ and the sample variance is calculated $s^2 = \frac{\sum_i (x_i - \overline{x})^2}{n-1}$. With these two sample estimates, we can calculate a sample estimate of the CV:

$$\widehat{\theta} = \frac{s}{\overline{x}}$$

For each hub-to-spoke and spoke-to-hub route k , for a given carrier c , we can compute the CV, $\widehat{\theta}_{kc}$. For a each carrier table 1a and table 1b report the mean and standard deviation of $\widehat{\theta}_{kc}$:

$$mean = \widehat{\widehat{\theta}}_{kc} = \sum_{k=1}^{K} \widehat{\theta}_{kc}$$

$$std = \left(\sum_{k=1}^{K} (\widehat{\theta}_{kc} - \widehat{\widehat{\theta}}_{kc})^2 \right)^{1/2}$$

Delta Method

To make inferences about $\widehat{\theta}$, it is necessary to know something about its variance. Let

ϕ and $\widehat{\phi}$ denote the population and sample estimate of the variance of the CV. While the true variance ϕ is unknown, we can use the delta method to make a sample estimate $\widehat{\phi}$. Using the delta method, the sampling variance of the coefficient of variation has been shown[1] to be

$$\widehat{\phi} = \frac{\theta^2}{n}\left[\frac{\mu_4 - \mu_2^2}{4\mu_2^2} + \frac{\mu_2}{\mu_1'^2} - \frac{\mu_3}{\mu_2 \mu_1'} \right]$$

where μ_j' represents the jth moment of the population and μ_j represents the jth moment about the mean. Specifically, μ_2 is the population variance of X, μ_1' is the population mean of X, μ_3 is the skewness and μ_4 is the kurtosis. Again, we do not know or observe the population levels of μ_1', μ_2, μ_3, or μ_4. However we can use the unbiased sample estimators:

$$\mu_j' \simeq \overline{x^j} = \frac{\sum_i x_i^j}{n}$$

$$\mu_2 \simeq s^2 = \frac{\sum_i (x_i - \overline{x})^2}{n-1} = \frac{n(\overline{x^2} - \overline{x}^2)}{n-1}$$

$$\mu_3 \simeq \frac{n^2(\overline{x^3} - 3\overline{x^2}\,\overline{x} + 2\overline{x}^3)}{(n-1)(n-2)}$$

$$\mu_4 \simeq \frac{n^2(n+1)(\overline{x^4} - 3\overline{x^3}\,\overline{x} + 6\overline{x^2}\,\overline{x}^2 - 3\overline{x}^4)}{(n-1)(n-2)(n-3)}$$

Thus, for direction $k \in \{1 = \text{hub-to-spoke}, 2 = \text{spoke-to-hub}\}$ we estimate $\widehat{\theta}_k$ and $\widehat{\phi}_k$. Armed with these sample estimates we can test various null and alternative hypotheses concerning the relative directional CV values. For instance:

$$H_o : \theta_1 - \theta_2 = 0$$

$$H_A : \theta_1 - \theta_2 \neq 0$$

Because of the central limit theorem, the sampling distribution of $\widehat{\theta}_1 - \widehat{\theta}_2$ can be approximated by a normal probability distribution for a large sample. Thus we compute the following statistic:

$$z = \frac{(\widehat{\theta}_1 - \widehat{\theta}_2) - (\theta_1 - \theta_2)}{\sqrt{(\phi_1/n_1) + (\phi_2/n_2)}}$$

Since ϕ_1 and ϕ_2 are unknown, we use the sample variances $\widehat{\phi_1}$ and $\widehat{\phi_2}$ to compute

[1]Kendall, Maurice G. and Stuart, Allan, *The Advanced Theory of Statistics*, New York: Macmillan Publishing, 1977

the test statistic.

Table 3 reports the results of the various hypothesis tests using the standard errors calculated using the delta method. Given $k \in \{1 = \text{hub-to-spoke}, 2 = \text{spoke-to-hub}\}$, we first test

$$H_0^I : \theta_1 - \theta_2 \leq 0$$
$$H_A^I : \theta_1 - \theta_2 > 0$$

Column three reports the number of airport pairs, for each carrier, for which we reject null-hypothesis H_0^I. In these airport-pairs, hub-spoke CV appears significantly higher than the spoke-hub CV.

An alternative one-sided test is the hypothesis:

$$H_0^{II} : \theta_1 - \theta_2 \geq 0$$
$$H_A^{II} : \theta_1 - \theta_2 < 0$$

Column four reports the number of airport pairs, for each carrier, for which we reject null-hypothesis H_0^{II}. Column five are the number of markets where we can neither reject H_0^I nor H_0^{II}.

<u>Bootstrapping Method</u>

Bootstrapping offers a second method for calculating standard errors. This technique involves studying a sample of bootstrap estimators $\widehat{\theta}(b)_m$, b = 1,...,B obtained by sampling m observations, with replacement, from X and recomputing $\widehat{\theta}$ with each sample. This is done a total of B times with the desired sampling characteristic computed from $\Theta = [\widehat{\theta}(1)_m, \ldots, \widehat{\theta}(B)_m]$. Because the estimator is consistent and n are reasonably large in each sample, we can approximate the asymptotic covariance matrix of the estimator $\widehat{\theta}$ by using:

$$\text{Est. Asy Var}[\hat{\theta}] = \frac{1}{B}\sum_{b=1}^{B}[\hat{\theta}(b)_m - \hat{\theta}][\hat{\theta}(b)_m - \hat{\theta}]'$$

We compute the estimated asymptotic variance using B=1000 samples of m=1000 observations for each unique hub-to-spoke and spoke-to-hub airport-pair-market.

Table 4 reports the results of the various hypothesis tests. Given $k \in \{1 = \text{hub-to-spoke}, 2 = \text{spoke-to-hub}\}$, We first test

$$H_0^I : \theta_1 - \theta_2 \leq 0$$
$$H_A^I : \theta_1 - \theta_2 > 0$$

Column three reports the number of airport pairs, for each carrier, for which we reject null-hypothesis H_0^I. In these airport-pairs, hub-spoke CV appears significantly higher than the spoke-hub CV.

An alternative one-sided test is the hypothesis:

$$H_0^{II}: \theta_1 - \theta_2 \geq 0$$
$$H_A^{II}: \theta_1 - \theta_2 < 0$$

Column four reports the number of airport pairs, for each carrier, for which we reject null-hypothesis H_0^{II}. Column five are the number of markets where we can neither reject H_0^{I} nor H_0^{II}.