ECONOMIC ANALYSIS GROUP DISCUSSION PAPER

Common Ownership and Airlines: Evaluating an Alternate Ownership Data Source.*

Eric Lewis† Randy Chugh

April 2019

EAG 19-1

---

* The views expressed herein are entirely those of the authors and should not be purported to reflect those of the US Department of Justice.

† US Department of Justice, Antitrust Division, Economic Analysis Group, 450 5th St. NW, Washington DC 20530. Email: Eric.Lewis@usdoj.gov

Common Ownership and Airlines: Evaluating an Alternate Ownership Data Source.

Eric Lewis, Randy Chugh[1]

April 17, 2019

Working Paper

Abstract:

Starting with the pioneering work of Azar, Schmalz, and Tecu (2018), a rapidly growing body of empirical evidence on the effect of common ownership on market outcomes has emerged. However, testing the robustness of these results to alternative methods and data sources is just beginning. In this paper, we contribute to this growing body of work by comparing results based on two different data sources on institutional ownership: Thomson Reuters Spectrum ("SP") and Thomson Reuters Ownership ("OP"). While SP is used by most researchers in this field, we find that OP has several distinct advantages including broader coverage and more convenient data formatting. We replicate the results of Azar, Schmalz, and Tecu (2018) and show that empirical results change dramatically when using OP instead of SP. We also find evidence that MHHI delta measures using OP data are more volatile than those using the SP data.

Introduction:

Institutional investors often own shares in multiple firms, including firms that compete against each other. Salop and O'Brien (2000) show that if firms maximize the total profits of their shareholders, and shareholders own stakes in multiple competing firms, then firms will have an incentive to reduce quantities and increase price above the Cournot equilibrium outcome. Indeed, if all firms have identical "portfolios" of owners, Salop and O'Brien show that firms will act as if all firms are run by a single monopolist. However, some factors may make that outcome difficult to achieve, including regulatory

[1] The authors are grateful to Andrew De Jong for excellent research assistance. The analysis and conclusions expressed herein are solely those of the authors and are not purported to represent the views of the United States Department of Justice.

scrutiny and principal-agent frictions. Therefore, how much common ownership reduces competition and raises prices is an empirical question.

One of the first forays into answering this question was by Azar, Schmalz, and Tecu (2018, hereafter "AST") who examine the effect of common ownership on prices using data on the US airline industry. For each route by quarter observation, they constructed a measure of concentration that takes into account common ownership. This measure is called the Modified Herfindahl-Hirschman Index (MHHI), and was first proposed by Salop and O'Brien (2000). They also construct the Herfindahl-Hirschman Index (HHI) which is the conventional measure of concentration.[2] AST regress route prices on both HHI and the difference between MHHI and HHI, which they call MHHI delta. The estimated coefficients on both HHI and MHHI delta are positive, statistically significant, and similar in magnitude, implying that increased concentration caused by institutional ownership has the same effect on prices as conventional concentration.

A number of additional papers have appeared that have examined common ownership in airlines and banking. Azar, Raina and Schmalz (2016) take a similar empirical approach, except that they examine the banking industry. Some papers have criticized both this paper and AST because the empirical specifications include right-hand-side variables like HHI and MHHI delta which are functions of endogenous market shares (O'Brien and Waehrer 2017, Rock and Rubinfeld 2017). To resolve the endogeneity issue, Kennedy et al. (2017) and Gramlich and Grundl (2017) use alternate empirical specifications that do not rely on regressing price on endogenous HHI and MHHI delta. In addition, Kennedy et al. (2017) estimate a structural model of common ownership where common ownership is incorporated into the firm's first order condition. Dennis, Gerardi and Schenone (2017) argue that AST's results are not robust to changes to sample selection and weighting.

In this paper, our focus is not on methodology but rather on the source of common ownership data. We discuss two different data sources that researchers may use to measure common ownership and MHHI. The first is the Thomson Reuters SP data product which is the one used by AST, Kennedy et al. (2017), and Dennis, Gerardi and Schenone (2017). The second is the Thomson Reuters OP product,

[2] In a conventional homogenous good Cournot model with no common ownership, HHI will be proportional to average margins. MHHI is an extension to this which takes into account common ownership: Salop and O'Brien show that if firms compete à la Cournot and maximize the total profits of their owners, the average margin will be proportional to the MHHI.

which to our knowledge has not previously been used in research on common ownership. We discuss the pros and cons of each data source.

Using the same airline setting as AST, we examine how their results change if MHHI delta is computed using OP data rather than SP data. We find that the coefficient on MHHI delta drops significantly when we use the OP data and, in many cases, loses statistical significance. We also find that the OP data seems to be measuring similar ownership as the SP data, as measures of MHHI delta from the two ownership data sources are strongly correlated and have similar time trends. Finally, we find that measures of MHHI delta using OP data exhibit more volatility than those using SP data.

Our results show that common ownership results are sensitive to ownership data sources. Further work should be done to better understand the differences between SP data, OP data, and other sources of ownership data (such as the raw 13F data). While some work on understanding ownership data has already done (for example Backus, Conlon, and Sinkinson, 2019), the fact that the 13F data (and therefore the SP data) suggest scope for a dataset like the OP data, although the proprietary construction methodology used to construct OP data has its own drawbacks.

In this paper we first discuss background information, including the SP and OP data products. Next, we discuss data set construction. Then we turn to our empirical results, comparing how results change when we use the OP data rather than the SP data.

Background:

The major source of institutional ownership data in the United States is SEC 13F filings. Investors with holdings over $100 million are required to report their shareholdings to the SEC as 13F filings, listing all firms that the investor owns shares in, the value of those shares, and their voting shares. The 13F filings are publicly available on the SEC's online database. However, 13F filings are challenging to use because they are not formatted uniformly[3] and they have reporting errors and other typos. In addition, 13F filings are not a complete picture of investors because they do not include investors with smaller holdings, individual investors, and foreign entities who own stock in US firms.

Thomson Reuters provides a cleaned version of 13F filings known as the Spectrum or SP data. SP is similar in format to the raw 13F filings in that it is organized by investor, where for each investor it lists all the firms in which investor owns stock, including stock holdings, value and voting share. The SP data

[3] Beginning in mid-2013, 13F forms were required to be reported in .xml format, which increased the ease of reading in the documents and reduced the rate of reporting errors (Backus, Conlon and Sinkinson 2019).

cleans many of the typos found in the raw 13F data. In addition to equity (stock holdings), SP data also includes fixed income (mostly preferred stock).

Thomson Reuters provides another data product known as the Ownership Product or OP data. Unlike SP which draws almost exclusively on SEC filings, OP draws on a wider variety of sources, including other regulatory bodies, stock exchanges, and fund reports received from fund management companies, etc., in addition to SEC filings. Therefore, it includes additional investors, including foreign investors and individuals, that are not in the SP and 13F data.[4] One very useful feature of OP is that it in the case of an investor who reports multiple 13F filings, the OP data attempts to aggregate shares to the investor level. Another very useful feature of the OP data is that it is organized by firm: For each firm, the data lists the "portfolio" of all owners along with their shares in the firm. This is particularly helpful for researchers studying the effects of common ownership within an industry as it requires them to only download data from the firms in that industry. One major drawback of the OP data is that it does not include voting shares.

Data set construction:

To replicate AST's results, we use their replication package provided on the paper's Journal of Finance web page.[5] Their replication package includes both data and replication code. Auxiliary data files, such as city populations and distances, are included. The large Department of Transportation data are not included but are easily downloaded online. The SP data are not included due to a data license restriction. The replication package does, however, include the MHHI delta measures that AST calculate from the SP data.

We find that it is a simple process to use AST's replication package to recreate AST's baseline data set. We refer to this build of the data as the AST replication rebuild.

We download OP ownership data for US airlines from the ThomsonOne.com website.[6] As discussed above, these data are at the airline level rather than at the investor level, and each airline by

[4] For example, Lawrence Page and Sergey Brin, the two individuals with the largest shares of Alphabet Inc. (Google), are listed in the OP data but not the SP data.
[5] https://onlinelibrary.wiley.com/doi/full/10.1111/jofi.12698
[6] Accessing ThomsonOne.com and the OP data requires having a paid subscription. From there, to find and download the share ownership data, log in to ThomsonOne.com. Click on "Screening & Analysis" in the lower left corner. From the menu of options that appears in the upper left corner, click "Share Ownership." According to Thomson Reuters representatives, ThomsonOne.com will soon be retired and the OP data products will be transferred to Thomson Reuter's Eikon platform.

quarter download is a list of the portfolio of owners along with their shares. We find that the airline identifier names change over time due to mergers, bankruptcy, and other events, and so we construct consistent airline identifiers such that the OP data can be merged with the DOT data.

While OP data does some work to aggregate filings in case a single investor files multiple 13F fillings, we find that investor names need to be further cleaned to reflect actual ownership, mapping the investor to its parent owner. We also find that some investors, especially the largest ones, sometimes still have multiple entities listed (e.g., "Blackrock (Netherlands) B.V." and "BlackRock Japan Co., Ltd"). Therefore, we take a number of approaches to clean the data: One method we use is to use publicly available information to map investors to their parent investors or actual owners. Another method we use is to only use the investor's first name as the unique identifier of the investor. This results in four total cleaning approaches which are summarized in Table 1.

Table 1: Summary of the various TR1 measures of MHHI delta and their correlation with the AST measure of MHHI delta. Versions 1 and 2 use publicly available information to map investors to the actual parent. Approaches 2 and 4 use only the first word of the name to identify investor.

Version of MHHI delta	Uses parent	Uses only first word of name	Correlation with AST MHHI delta
1	Yes	No	0.755
2	Yes	Yes	0.777
3	No	No	0.706
4	No	Yes	0.755

To construct airline-specific variables like average price and market shares, we do a separate build of the airline data which we refer to as the OAG build. We use OAG data rather than DOT's T-100 data (as in AST) because it corrects for cases where the DOT data reports the regional carrier but the flight is sold by the ticketing carrier (see Aryal, Ciliberto, and Leyden 2019). In all other respects we follow AST's instructions in building the airline data. Due to some licensing restrictions on OAG data, our OAG build begins in 2002, whereas AST's begins in 2001. We find that our sample of markets almost perfectly matches AST's sample: There are only three infrequently traveled routes that appear in the AST data but not in our OAG build.

When comparing AST's replication rebuild with our OAG build, we find that the build is almost identical. Within markets, log fares and HHIs match up almost exactly, with correlations exceeding 99.95%.[7] This implies that the OAG data does not significantly shift shares, and that AST's measure of

[7] Correlation of log fare between the builds is 0.9999. Correlation of HHIs between the two builds is 0.9996.

MHHI delta would not change significantly if they had used OAG data. To calculate the OP versions of MHHI and MHHI delta, we use the OAG build as the source of shares, creating four different versions of MHHI delta which depend on how we clean the OP data, as discussed above and shown in Table 1.

Findings:

We now turn to our empirical results. Throughout this paper we focus on the market-level (e.g., route by quarter) results rather than the airline by market-level results. We do this because the variation in HHI and MHHI delta happens at the market level. Our findings are broadly similar if our observations are at the airline by market level.

We first examine how well regressions using our AST replication build match that with the results reported in the published AST paper. Our replication of AST's baseline regression results are in Table 2 columns 1-3. (The corresponding original AST results are presented in AST's Table 3 columns 4-6, available at the Journal of Finance website). We find that these coefficients are very similar, though minor discrepancies exist. For example, one discrepancy is seen in the first specification (column 1 of our Table 2 and column 4 of Table 3 in AST's paper): AST's original results find the coefficient on MHHI delta is 0.325 whereas we find that the coefficient is 0.321. We find that these differences are small and likely due either to rounding error or updates to the DOT data.

Next, we examine how the AST rebuild results would change if we limit the AST rebuild sample to those market by quarter observations which are in the OAG build. We do this comparison in order to ensure that AST results are robust to small changes in sample. As discussed above, the major change in sample is that year 2001 observations are excluded from the OAG build. When making this restriction, we find that AST's findings continue to hold: The coefficient on MHHI delta remains large and significant and nearly the same as the coefficient on HHI. Therefore, we conclude that changing the sample to remove 2001 does not have a significant effect on AST's baseline results. This comparison can be seen by comparing columns 4-6 of Table 2 (restricted sample) with columns 1-3 of Table 2 (full AST sample).

Table 2: *Regression results where we compare AST's sample size in columns 1-3 with a restricted sample in columns 4-6. Columns 4-6 exclude 2001 as well as 3 observations that were not in the OAG build.*

VARIABLES	(1) log(fare)	(2) log(fare)	(3) log(fare)	(4) log(fare)	(5) log(fare)	(6) log(fare)
MHHI delta	0.321***	0.314***	0.203***	0.289***	0.298***	0.170***
	(0.0447)	(0.0396)	(0.0353)	(0.0469)	(0.0430)	(0.0339)
HHI	0.365***	0.359***	0.259***	0.341***	0.338***	0.236***
	(0.0317)	(0.0314)	(0.0242)	(0.0309)	(0.0310)	(0.0226)
Number of Nonstop Carriers			-0.00822**			-0.00475
			(0.00352)			(0.00334)
Southwest Indicator			-0.150***			-0.149***
			(0.0135)			(0.0129)
Other LCC Indicator			-0.100***			-0.105***
			(0.00993)			(0.00963)
Share of Passengers Traveling Connect, Market-Level			0.176***			0.180***
			(0.0193)			(0.0191)
Share of Passengers Traveling Connect			0			0
			(0)			(6.16e-09)
Log(Population)			0.350***			0.185
			(0.122)			(0.122)
Log(Income Per Capita)			0.314***			0.198*
			(0.109)			(0.100)
Constant	4.932***	0.479	3.269***	4.945***	0.0903	3.709***
	(0.0221)	(0.708)	(0.694)	(0.0219)	(0.713)	(0.693)
Observations	262,308	262,308	254,957	244,179	244,179	237,363
R-squared	0.853	0.861	0.876	0.859	0.866	0.881
market FE	Yes	Yes	Yes	Yes	Yes	Yes
log distance x market FE	No	Yes	Yes	No	Yes	Yes
Limited to OAG build	No	No	No	Yes	Yes	Yes
Number of markets	7184	7184	6905	7077	7077	6809

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Next, we examine how closely our four versions of the OP MHHI delta compare to the SP MHHI delta calculated by AST. We find that the correlation between the SP MHHI delta and the various OP MHHI delta measures ranges from 70.6% to 77.7% (see Table 2). We find that using the parent rather than the filer name (versions 1 and 2) tends to increase the correlation of the OP MHHI delta with the SP MHHI delta. Similarly, using the first word of the cleaned investor name (versions 2 and 4) also increases

the correlation with AST's MHHI delta. The most strongly correlated measure of OP MHHI delta is where we do both types of cleaning (version 2).

We also examine how the within-market correlation of SP MHHI delta and the four measures of OP MHHI delta changes over time. Figure 1 shows that the correlation between these measures ranges from about 40% to more than 90%. This suggests that while the two ownership sources of data are measuring similar investors, there are periods where the two do not match up well, which may be driven by different inputs or by changing data construction methodologies.

In Figure 2, we plot the average MHHI delta using these five different measures. Overall, the SP MHHI delta has similar trends to the four measures of OP MHHI delta, but the actual values differ significantly. All exhibit declines in 2005 as well as a long run overall increase starting in about 2007. The OP measures of MHHI delta exhibit overall more volatility and a greater range than the SP MHHI delta. Version 3 of the OP MHHI delta (which has no additional cleaning) is significantly different from the other three versions of the OP MHHI delta, and is always lower in value than the SP MHHI delta. Again, this suggests significant differences in either the data sources or the data construction.

Finally, we turn to regression results, and examine how the baseline regression specification results change if we use the various OP measures of MHHI delta rather than the SP MHHI delta. To do this, we merge the OP MHHI delta measures into the AST replication rebuild dataset. Regression results are in Tables 3, 4, and 5. Table 3 contains results where we include only market fixed effects and is the same specification as columns 1 and 4 of Table 2. Table 4 adds distance interacted with market fixed effects and is the same specification as columns 2 and 5 of Table 2. Table 5 adds in additional controls and is the same specification as columns 3 and 6 of Table 2.

For all results in Tables 3, 4, and 5, we find that when using the various versions of the OP MHHI delta, the coefficient on MHHI delta drops to close to zero for all specifications other than the one that uses version 3 of the OP MHHI delta. In fact, for versions 1, 2, and 4, the point estimate is usually negative. For version 3, which takes the filer name as given, we find that the coefficient on MHHI delta is still significantly less than that on HHI, and we reject the null hypothesis that the two are equal.

Figure 1: Correlation between OP measures of MHHI delta and the SP MHHI delta.

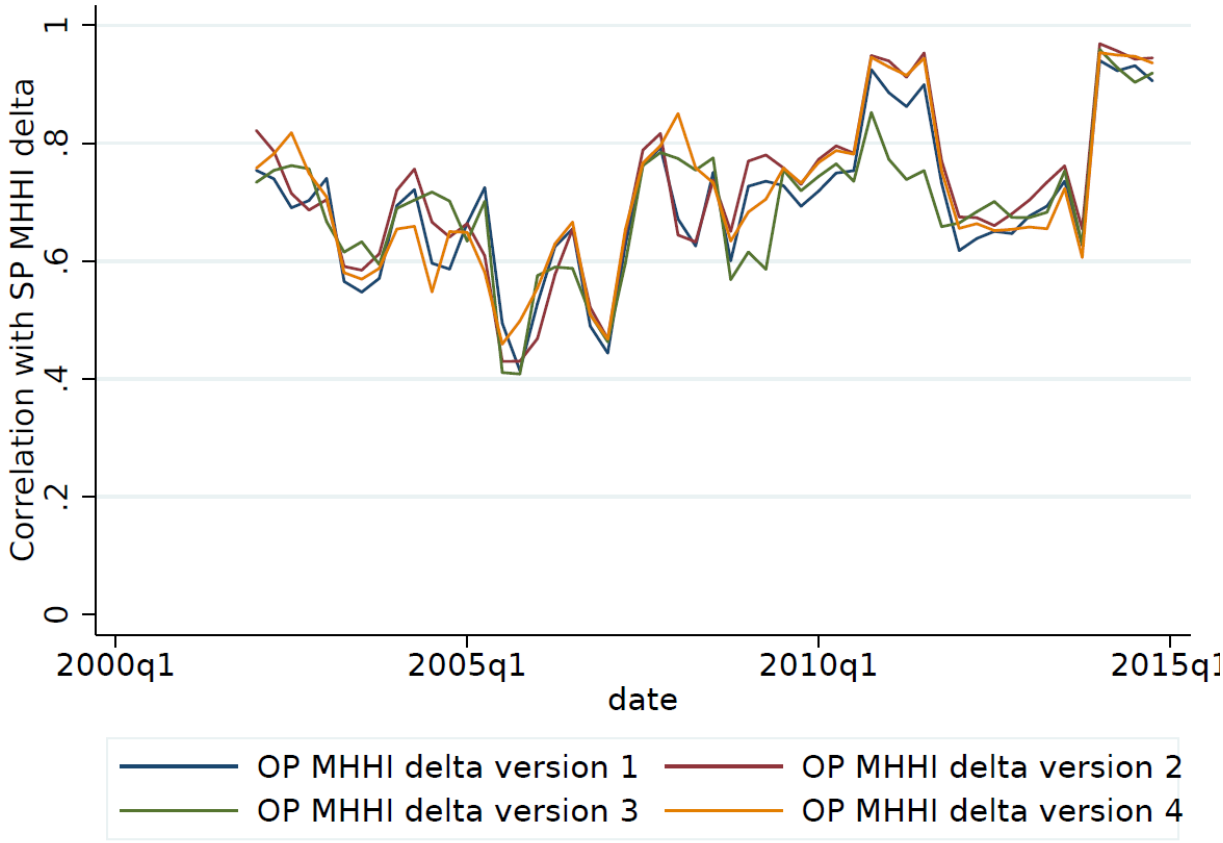

Figure 2: Average MHHI delta across routes within quarter.

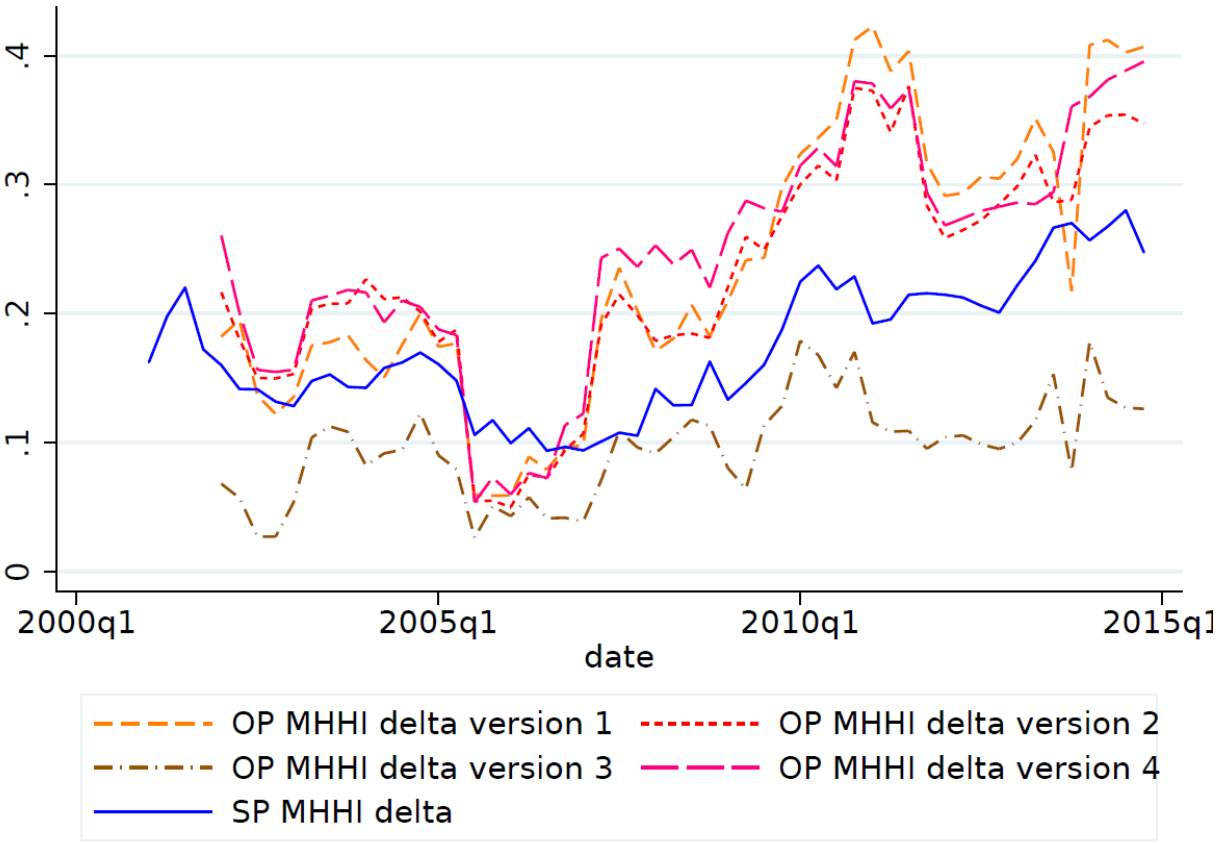

Table 1: Regressions using OP versions of MHHI delta. Regressions controls include market fixed effects. These use the same specification as AST's Table 3 column 4 and our Table2 columns 1 and 4.

VARIABLES	(1) log(fare)	(2) log(fare)	(3) log(fare)	(4) log(fare)
MHHI delta	-0.0265	-0.0233	0.155***	0.0149
	(0.0246)	(0.0245)	(0.0528)	(0.0242)
HHI	0.255***	0.257***	0.290***	0.272***
	(0.0296)	(0.0295)	(0.0283)	(0.0296)
Constant	5.040***	5.038***	5.002***	5.021***
	(0.0198)	(0.0196)	(0.0183)	(0.0199)
Observations	244,179	244,179	244,179	244,179
R-squared	0.857	0.857	0.857	0.857
market FE	Yes	Yes	Yes	Yes
log distance x market FE	No	No	No	No
OP MHHI delta version:	1	2	3	4
Number of markets	7077	7077	7077	7077

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2: Regressions using OP versions of MHHI delta. Regressions controls include market fixed effects and distance interacted with market fixed effects. These use the same specification as AST's Table 3 column 5 and our Table 2 columns 2 and 5.

VARIABLES	(1) log(fare)	(2) log(fare)	(3) log(fare)	(4) log(fare)
MHHI delta	-0.00265	-0.0120	0.173***	0.0230
	(0.0241)	(0.0237)	(0.0480)	(0.0233)
HHI	0.262***	0.258***	0.289***	0.272***
	(0.0298)	(0.0296)	(0.0287)	(0.0299)
Constant	-0.213	-0.219	-0.138	-0.200
	(0.721)	(0.720)	(0.718)	(0.722)
Observations	244,179	244,179	244,179	244,179
R-squared	0.864	0.864	0.865	0.864
market FE	Yes	Yes	Yes	Yes
log distance x market FE	Yes	Yes	Yes	Yes
OP MHHI delta version:	1	2	3	4
Number of markets	7077	7077	7077	7077

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3: Regressions using OP versions of MHHI delta. Regressions controls include market fixed effects, distance interacted with market fixed effects, and additional controls as shown below. These use the same specification as AST's Table 3 column 6 and our Table 2 columns 3 and 6.

VARIABLES	(1) log(fare)	(2) log(fare)	(3) log(fare)	(4) log(fare)
MHHI delta	-0.0147	-0.0270	0.0816**	-0.00346
	(0.0186)	(0.0181)	(0.0381)	(0.0169)
HHI	0.182***	0.177***	0.203***	0.187***
	(0.0221)	(0.0221)	(0.0214)	(0.0223)
Number of Nonstop Carriers	-0.00570*	-0.00593*	-0.00491	-0.00548
	(0.00338)	(0.00337)	(0.00345)	(0.00338)
Southwest Indicator	-0.147***	-0.146***	-0.148***	-0.148***
	(0.0132)	(0.0131)	(0.0131)	(0.0132)
Other LCC Indicator	-0.113***	-0.113***	-0.112***	-0.112***
	(0.0102)	(0.0102)	(0.0101)	(0.0102)
Share of Passengers Traveling Connect, Market-Level	0.186***	0.186***	0.184***	0.186***
	(0.0192)	(0.0192)	(0.0192)	(0.0191)
Share of Passengers Traveling Connect	0	0	0	0
	(0)	(0)	(1.34e-09)	(0)
Log(Population)	0.167	0.167	0.163	0.166
	(0.124)	(0.124)	(0.124)	(0.124)
Log(Income Per Capita)	0.195*	0.195*	0.198*	0.197*
	(0.100)	(0.100)	(0.101)	(0.100)
Constant	3.767***	3.766***	3.739***	3.749***
	(0.681)	(0.680)	(0.688)	(0.683)
Observations	237,363	237,363	237,363	237,363
R-squared	0.881	0.881	0.881	0.881
market FE	Yes	Yes	Yes	Yes
log distance x market FE	Yes	Yes	Yes	Yes
OP MHHI delta version:	1	2	3	4
Number of markets	6809	6809	6809	6809

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

It is unclear why these OP measures of MHHI delta give such different results than the SP MHHI delta. One possibility is that the SP data is not measuring all relevant components of institutional ownership, such as international investors, and therefore is misstating ownership. An alternative explanation is that the OP ownership data is not constructed in a consistent manner and therefore exhibits greater measurement error and volatility. Future research should examine why the two data sources generate such different results.

Conclusion:

We find evidence that AST's results on common ownership are not robust to source of ownership data. In particular, we cannot replicate AST's results if we use Thomson Reuter's Ownership Product (OP) data as the ownership data source rather than Thomson Reuter's Spectrum (SP) data. We find that the coefficient on MHHI delta is much closer to zero, statistically insignificant, and usually negative when we use the OP data rather than the SP data.

Future research should work to understand why these two ownership datasets give such different results and to create a new ownership data that combines the strengths of the OP and the SP data. Important attributes of such a data set are consistent cleaning and construction, use of additional data sources beyond the SEC 13F, and data that can be queried by firm (as opposed to investor). Progress in understanding and cleaning institutional ownership data can help researchers and policy makers better understand any effects of institutional ownership on competition. Significant work in this area has already been done by Backus, Conlon and Sinkinson (2019) who provide resources to clean and analyze the publicly available 13F data, and AST, who provide MHHI delta calculations for air travel markets from the SP data.

References:

Aryal, Gaurab, Federico Ciliberto, and Benjamin T. Leyden (2019). "Coordinated Capacity Reductions and Public Communication in the Airline Industry". Working paper.

Azar, Jose, Martin C. Schmalz, and Isabel Tecu (2018) "Anticompetitive Effects of Common Ownership." *The Journal of Finance.* 73(4).

Azar, Jose, Sahil Raina, and Martin Schmalz (2016) "Ultimate Ownership and Bank Competition". Working paper.

Backus, Matthew, Christopher Conlon, and Michael Sinkinson (2019). "Common Ownership in America: 1980-2017." Working Paper.

Dennis, Patrick, Kristopher Gerardi, and Carola Schenone (2017) "Common Ownership Does Not Have Anti-Competitive Effects in the Airline Industry." Working Paper.

Elhauge, Einer (2016) "Horizontal Shareholding". *Harvard Law Review*, 129.

Gramlich, Jacob and Serafin Grundl (2017) "Testing for Competitive Effects of Common Ownership," Finance and Economics Discussion Series 2017-029. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2017.029.

Kennedy, Pauline, Daniel P. O'Brien, Minjae Song, and Keith Waehrer. (2017). "The Competitive Effects of Common Ownership: Economic Foundations and Empirical Evidence." Working Paper.

O'Brien, Daniel P. and Keith Waherer (2017): "The Competitive Effects of Common Ownership: We Know Less Than We Think." Working Paper.

Posner, Eric A., Fiona Scott Morton, and E. Glen Weyl (2017). "A Proposal to Limit the Anti-Competitive Power of Institutional Investors." Working Paper.

Rock, Edward B. and Daniel L. Rubinfeld (2017). "Defusing the Antitrust Threat to Institutional Investor Involvement in Corporate Governance." Working Paper.

Salop, Steven C. and Daniel P. O'Brien (2000). "Competitive Effects of Partial Ownership: Financial Interest and Corporate Control." *Antitrust Law Journal.* 67, pp 559-614.