THE JOURNAL OF INDUSTRIAL ECONOMICS 0022-1821
Volume LX September 2012 No. 3

PRICE DISPERSION OVER THE BUSINESS CYCLE: EVIDENCE FROM THE AIRLINE INDUSTRY*

Marco Cornia[†]
Kristopher S. Gerardi[‡]
Adam Hale Shapiro[§]

This study provides empirical evidence documenting how price dispersion moves with the business cycle in the airline industry. Performing a fixed-effects panel analysis on seventeen years of data covering two business cycles, we find that price dispersion is highly pro-cyclical. This effect is especially pronounced for legacy carriers relative to low-cost carriers. We show that our empirical result is consistent with firms' implementing second-degree price-discrimination tactics.

I. INTRODUCTION

Economists have long been captivated by the fact that for many homogenous goods, a distribution of prices exists rather than a single price. Numerous empirical studies (for example, Shepard [1991], Sorensen [2000], Stavins [2001], and Hendel and Nevo [2011]) and theoretical models (for example, Salop and Stiglitz [1982], Burdett and Judd [1983], Holmes [1989], and Aguirre, Cowan and Vickers [2010]) have been produced to better understand this phenomenon, but its fundamental causes are still widely debated. This paper adds to the empirical literature on this topic by providing evidence of how price dispersion moves with the business cycle.[1] Understanding how aggregate factors affect price dispersion may ultimately provide economists with a better understanding of firms' pricing decisions.

*We thank Jim Adams, Ana Aizcorbe, Ben Bridgman, Abe Dunn, Matt Grennan, Matt Osborne, and Kyle Hood for helpful comments and discussions. We also thank the Editor, as well as two anonymous referees for suggestions that greatly improved the paper. The views expressed in this paper are solely those of the authors and do not necessarily reflect the views of the Bureau of Economic Analysis, the U.S. Department of Commerce, the Federal Reserve Bank of Atlanta or the Federal Reserve System. The research presented here was primarily conducted while at the Federal Reserve Bank of Boston.

[†]Authors' affiliations: Graduate Student, Johns Hopkins University, Baltimore, Maryland, U.S.A.
e-mail: mcornia1@jhu.edu
[‡]Federal Reserve Bank of Atlanta, Atlanta, Georgia, U.S.A.
e-mail: kristopher.gerardi@atl.frb.org
[§]Federal Reserve Bank of San Francisco, San Francisco, California, U.S.A.
e-mail: adam.shapiro@sf.frb.org

[1] Additional studies on the determinants of price dispersion include Baye, Morgan and Scholten [2004], Goldberg and Verboven [2001], Gaggero and Piga [2011], and Orlov [2011].

The airline industry has been the focus of many empirical studies on price dispersion for a few important reasons. First, the airline industry is one in which firms are well known to charge a distribution of prices for the same product. Thus, there exists a large degree of price dispersion in the industry. Second, markets in the airline industry are cleanly delineated by distinct routes, which allows researchers to assess price dispersion empirically through panel-data methods. Finally, high quality data on airline prices and costs at relatively granular levels are publicly available.

In this study, we examine how various measures of price dispersion at the route level in the airline industry are correlated with the business cycle, while controlling for variation in price dispersion that is likely due to other factors, such as market structure, fuel, and cost variations. Our main result is that price dispersion moves pro-cyclically in the airline industry. Using a fixed-effects estimation on a panel that spans almost two full business cycles, we find that a rise in the output gap—a measure of the difference between nominal GDP and 'potential' GDP as defined by the Congressional Budget Office—of 1 percentage point is associated with a 1.5 per cent increase in the interquartile range, on average. A fall in the average city-endpoint unemployment rate of one percentage point causes a 2.3 per cent rise in the interquartile range, on average. Our results are robust to a range of different measures of price dispersion, including the Gini coefficient. Previous studies have found it important to differentiate between legacy carriers, also known as 'hub-and-spoke' carriers, and low-cost carriers (LCC's), because they behave quite differently along dimensions related to pricing, competition and network formation. Interestingly, our results indicate that price dispersion is more pro-cyclical for legacy carriers than it is for LCC's.

There are a number of potential mechanisms that could cause price dispersion to move pro-cyclically. With the available data on prices, costs and purchaser demographics, we are unable to single out any one particular mechanism as the sole contributor to this empirical finding. For instance, since we do not observe many of the individual ticket characteristics we cannot rule out the possibility that price dispersion varies due to a change in the tickets consumers purchase. However, we are able to provide theoretical and empirical evidence that favors some explanations over others. In particular, we provide evidence that pro-cyclical price dispersion may be a simple outcome of second-degree price discrimination tactics. We also provide empirical evidence that downplays the importance of stochastic-demand pricing (Eden [1990]). This result corresponds well with the recent findings of Puller, Sengupta and Wiggins [2009] who find that airline price dispersion is driven primarily by second-degree price discrimination tactics, as opposed to stochastic-demand pricing techniques.

This study is related to the growing literature on price dispersion in the airline industry. It is also related to numerous microeconomic studies on pricing strategies and business cycle conditions. For instance, Rotemberg and Saloner [1986] theorize that during booms, firms may be less likely to collude since the benefits of cheating are higher, causing firms to cut prices. There are a number of empirical papers on this topic that document The fact that retail prices tend to fall during periods of peak demand (see Warner and Barsky [1995], MacDonald [2000], Chevalier, Kashyap and Rossi [2003], and Nevo and Hatzitaskos [2006]). Another set of theories, based on switching costs and brand loyalty, show that during booms new customers may enter the market causing demand to become more elastic and firms to lower prices (see Bils [1989], Klemperer [1995], and Stiglitz [1984]). A third theory, put forth by Greenwald, Stiglitz and Weiss [1984] and analyzed by Chevalier and Scharfstein [1996], shows that during recessions, cash-strapped firms may forego offering low prices to attract new customers in order to generate a higher cash flow.

Although our empirical analysis is confined to one industry, we believe it likely has implications for other industries as well. If the correlation between airline price dispersion and measures of the business cycle that we document is due in part to price discrimination tactics, then we would expect to find pro-cyclical price dispersion in industries that are characterized by firms with market power and the ability to implement discriminatory pricing strategies such as hotels, stadiums, restaurants, theaters (Leslie [2004]), yellow-page advertising (Busse and Rysman [2005]), cement (Miller and Osborne [2010]) and personal computers (Aizcorbe and Shapiro [2010]).

The paper is structured as follows: Section 2 contains a detailed discussion of the data. In Section 3 we perform a fixed-effects, panel estimation of the relationship between price dispersion and various proxies for the business cycle. In Section 4 we provide a discussion of our empirical findings, paying particular attention to two leading theories of price dispersion: price discrimination and stochastic-demand pricing. We conclude in Section 5.

II. DATA

The empirical analysis focuses on domestic, direct, coach-class airline tickets over the period 1993q1 to 2009q4. The sample is constructed in the same manner as in Gerardi and Shapiro [2009] and includes nine major domestic airlines, often referred to as 'legacy carriers,'[2] as well as a number of low-cost carriers[3] (LCC's) and regional carriers. Ticket prices represent

[2] The legacy carriers in our sample include United, U.S. Airways, Delta, American, Alaskan, TWA, Continental, Northwest, and America West.

[3] The list of LCC's, obtained from Ito and Lee [2003], includes Air South, Access Air, AirTran, American Trans Air, Eastwind, Frontier, JetBlue, Kiwi, Morris Air, National, Pro Air, Reno, Southwest, Spirit, Sun Country, ValuJet, Vanguard, and Western Pacific. For a more detailed discussion of LCC's see Goolsbee and Syverson [2008].

TABLE I
SUMMARY STATISTICS

	Full Sample		Legacy		LCC		Big-City		Leisure	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Gini	0.22	0.08	0.25	0.07	0.17	0.05	0.25	0.08	0.19	0.07
IQR	92	83	112	92	52	27	117	105	64	51
90th Perc. Price	278	161	327	170	178	70	338	201	232	125
10th Perc. Price	95	39	102	40	80	29	99	37	98	48
HERF	0.76	0.25	0.75	0.25	0.77	0.25	0.67	0.24	0.77	0.25
COST	3.25	0.81	3.37	0.59	2.83	0.52	3.27	0.65	3.21	0.88
FUEL	1.26	0.98	1.15	0.82	1.45	0.76	1.24	0.84	1.41	1.36
UTIL	0.69	0.15	0.69	0.15	0.69	0.15	0.69	0.15	0.73	0.15

Notes: The interquartile range (IQR), 90th percentile price (90th Pctl.), 10th percentile price (10th Pctl.), and fuel cost per gallon (FUEL) are reported in dollars. Total operating fuel cost less fuel per seat-mile (COST) is reported in cents.

10 per cent of all domestic tickets issued by airlines and are obtained from the DB1B database. In addition to ticket prices, the DB1B includes other quarterly itinerary information, such as origin and destination airports, passenger quantities, number of stops (plane changes), and fare class.[4] Tickets less than 20 dollars are believed to be frequent-flyer tickets and are eliminated.

The data is a panel, where an observation is a flight conducted by a specific airline, between an origin and destination airport (route), in a specific time period (year and quarter). For example, an American Airlines direct, coach-class ticket, from Dallas (DFW) to San Francisco (SFO) in the first quarter of 1999 is considered an observation in our data. The direct ticket data include both one-way flights and round-trip flights. The DB1B contains numerous itineraries and fares for the same flight by the same carrier, reflecting the quarterly frequency of the data, as well as the many different fares found within the same fare class, on the same flight, at a given point in time. Thus, the data comprise distributions of prices for carrier-route itineraries.[5] Price dispersion is measured using three separate proxies: the interquartile range, the Gini coefficient, and the 90th and 10th price percentiles estimated separately. The interquartile range and Gini coefficient are advantageous in that they summarize dispersion with one statistic, while the price percentiles have the advantage that they provide more detailed information about the tails of the distribution.

Table I displays summary statistics of the variables that we include in our regression analysis. The mean Gini coefficient in our entire sample is 0.22,

[4] There are three different sub-components to the DB1B data set. They are market data, coupon data, and ticket data; and we combine variables from all three. For further reference, see the BTS's website http://www.transtats.bts.gov.

[5] See Appendix B for more details on the construction of the dataset, and Gerardi and Shapiro [2009] for an even more detailed description.

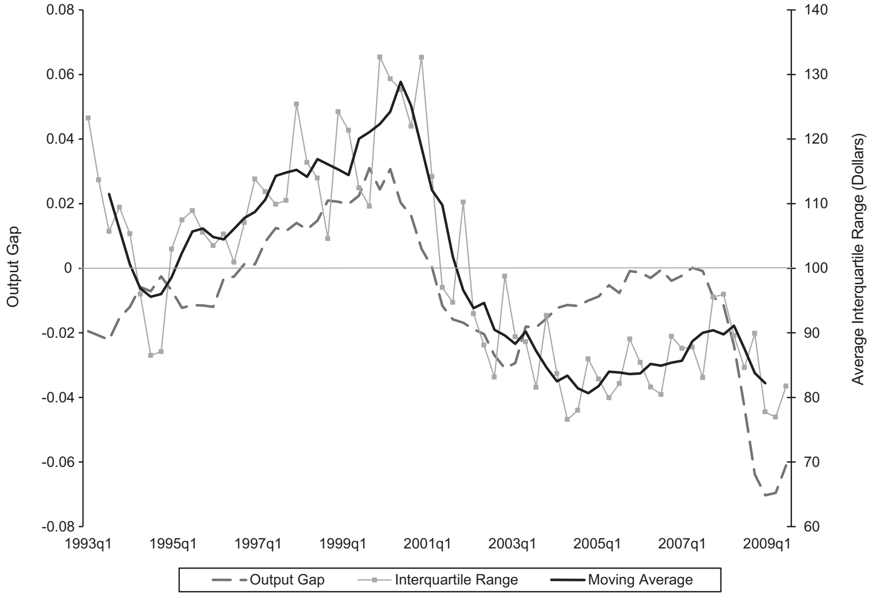

Figure 1
Price Dispersion over the Business Cycle

Notes: Depicted as a grey line is the passenger-weighted average of the interquartile range for all routes in the DB1B database. The solid black line is the five-quarter moving average. The output gap, as measured by the Congressional Budget Office (CBO), is also depicted as a dashed line.

and is 0.25 for legacy carriers and 0.17 for LCC's. The Gini coefficient can be shown to be equal to twice the expected absolute difference between two ticket prices drawn randomly from the population. For example, the median Gini coefficient for the entire data set is 0.225, which corresponds to an expected fare difference of 45 per cent of the mean fare for two randomly selected passengers. The mean interquartile range (IQR) is 92 dollars for our entire sample, and is 112 dollars for legacy carriers and 52 dollars for LCC's. Figure 1 plots the passenger-weighted average of the IQR over the sample period, along with the output gap, as measured by the Congressional Budget Office (CBO). The average degree of price dispersion rises with the boom in the late 1990's and then falls with the ensuing recession. Dispersion is flat throughout the mid 2000's, during which time the output gap was roughly zero, and then dispersion falls somewhat with the latest recession in 2008.

We present a few graphical examples of the pricing patterns seen in the data in order to show in more granular detail the dynamics of price dispersion. Figure 2 plots price percentiles of three routes along with a plot of the output gap. The output gap is defined as the log difference between the actual

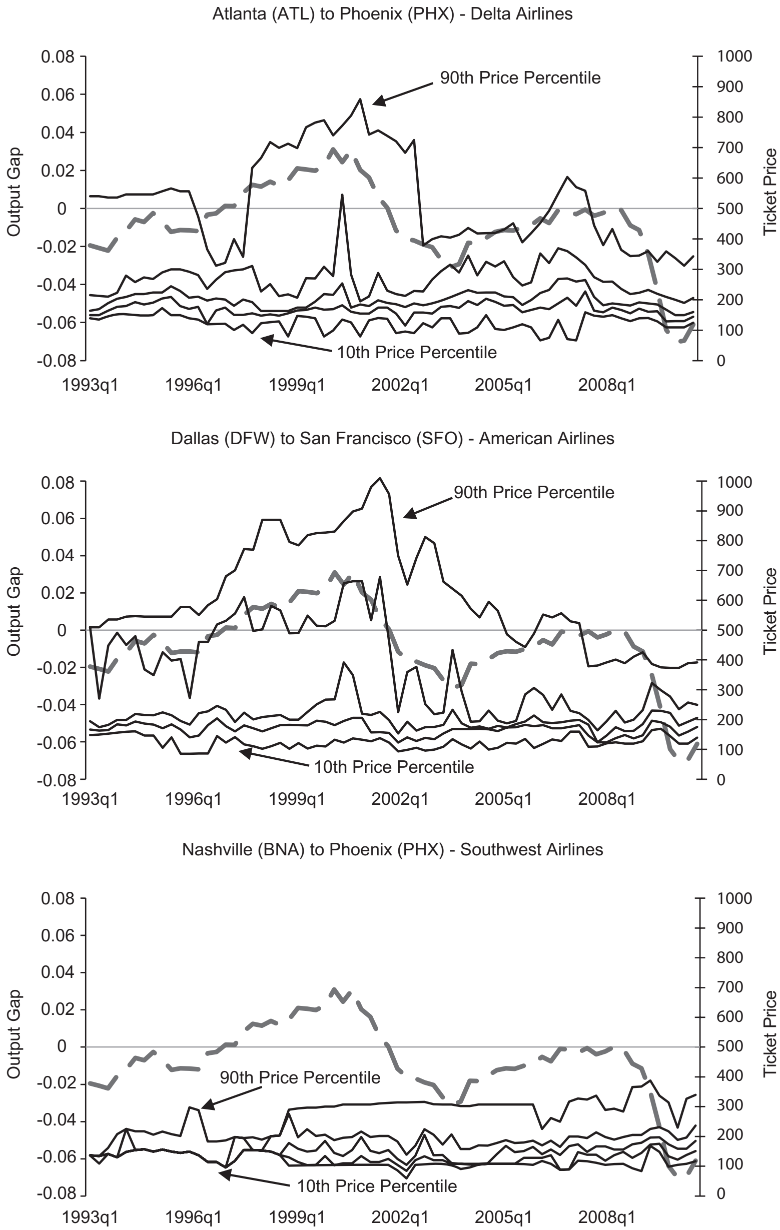

Figure 2
Pricing Dynamics

Notes: Depicted are 10th, 30th, 50th, 70th and 90th price percentiles for three airline-route observations. The output gap, as measured by the Congressional Budget Office (CBO), is also depicted as a dashed line.

nominal GDP and the CBO's measure of potential output. The top two panels correspond to routes operated by two legacy carriers, American Airlines and Delta Airlines, while the bottom panel consists of a route operated by Southwest Airlines. It is noteworthy that in the legacy carrier panels, the higher price percentiles seem closely to follow the output gap. The top portion of the price distribution rises and falls with the boom in the late 1990's and then begins gradually to fall as aggregate demand deteriorates. In contrast, we do not see the same relationship in the Southwest panel.

II(i). *Operating Cost*

As we are interested in studying variation in price dispersion that cannot be explained by variation in airline operating costs alone, we must include a control for the airlines' marginal cost in our empirical analysis. Airline marginal costs may vary over the business cycle for many reasons. For instance, wages of pilots and flight attendants may rise during booms, as may the price of fuel. We proxy for variations in marginal cost using a measure of the carrier's average variable cost. Numerous studies, such as Caves, Christensen and Tretheway [1984] and Gillen, Oum and Tretheway [1990], have found that the carriers' passenger output displays constant-returns-to-scale in firm size. This finding suggests that average variable cost may be a valid approximation to marginal cost in this context. We exploit the rich cost data available in the BTS P-52 database. Specifically, the BTS defines a measure called the 'total aircraft operating cost,' which includes fuel, crew wages, maintenance, aircraft leasing and depreciation. We are also able to decompose this variable into its fuel component and its other components. Due to the large market power of unions in the airline industry, non-fuel costs are particularly rigid relative to fuel costs.

Figure 3 plots total aircraft operating cost (including fuel) as a proportion of total seat-miles for four carriers over the sample period. The figure shows that cost per seat-mile is correlated across firms, and has generally increased through the course of the sample period. The large rise and fall in costs in 2008 can be attributed to the spike in oil prices that occurred during that summer. Southwest and JetBlue, the two largest LCC's in our sample, have lower cost levels relative to the two legacy carriers, U.S. Airways and United. This differentiation in cost between legacies and LCC's is ubiquitous across the entire airline industry. Table I provides summary statistics for our cost measures used in the empirical analysis. Total aircraft operating cost (less fuel) as a proportion of total seat-miles, *COST*, are higher on average for legacy carriers: 3.4 cents per seat-mile for legacy carriers as opposed to 2.8 cents per seat mile for LCC's. However, fuel costs (*FUEL*), measured as price per gallon, are higher for LCC's. Overall, including a proxy for marginal cost in the empirical specification removes any variation in price dispersion induced by variation in tangible costs.

Figure 3
Aircraft Operating Costs

Notes: Depicted are total aircraft operating cost (including fuel) in dollars per seat-mile for four carriers: U.S. Airways, United Airlines, JetBlue, and Southwest.

III. ESTIMATION

Since the data is a panel of airline-route observations, it is possible to assess the effects of business cycle variation on price dispersion while holding fixed time-invariant, route-specific factors, as well as any route-specific variation in the degree of competition and carrier-specific variation in fuel and other operating costs. We use a fixed-effects panel estimator, which exploits the time-series variation along a specific route in the estimation routine. We use two different approaches to measure the effect of business cycle variation on price dispersion.

The first specification takes the form:

$$\text{DISP}_{ijt} = \theta_0 + \beta * YGAP_t + \gamma_1 * \widehat{HERF}_{jt} + \gamma_2 \ln FUEL_{it} + \gamma_3 \ln COST_{it} + \delta_q + \nu_{ij} + \varepsilon_{ijt}. \tag{1}$$

where i indexes the carrier, j the route, t the specific time period, and q the quarter. In this specification, the output gap, $YGAP_t$ is used to proxy for the business cycle, as measured by the CBO, and carrier-route fixed effects are represented as ν_{ij}. We include the Herfindahl index, $\widehat{HERF}_{jt}$, to control for variation in market concentration of the route. As this measure is endogenous, we instrument using the same variables as in Borenstein and Rose [1994] and Gerardi and Shapiro [2009]. These instruments include the total number of enplaned passengers on the route, a measure of predicted concentration, and a measure of the airline's share of enplaned passengers at both endpoints. These variables are meant to capture exogenous variation in the degree of competition that are not directly correlated with the firm's pricing decision. We control for time-series variation in costs on a specific carrier i with the logarithm of the carrier's average fuel cost per gallon, ln $FUEL_{it}$, as well as the remaining operating cost per seat-mile, ln $COST_{it}$, measured by the BTS for a specific carrier. We also include quarter dummies, δ_q, to control for seasonal fluctuations.

The second specification takes the form:

$$\text{DISP}_{ijt} = \theta_0 + \beta * UR_{jt} + \gamma_1 * \widehat{HERF}_{jt} + \gamma_2 \ln FUEL_{it} + \gamma_3 \ln COST_{it} + \delta_q + \nu_{ij} + \varepsilon_{ijt}. \tag{2}$$

where the average unemployment rate of the two endpoint states on the route obtained from the Bureau of Labor Statistics, UR_{jt}, is used as an alternative proxy for the business cycle. In both specifications, price dispersion, $DISP_{ijt}$, is measured in three different ways: the logarithm of the interquartile range, the Gini log-odds ratio,[6] and the 90th and 10th percentiles, each estimated in separate regressions. Analyzing the top and bottom of the price distribution separately provides additional information regarding the source of the change in price dispersion. Observations are weighted by the total number of passengers on the route over the entire sample period and standard errors are clustered by route in order to control for autocorrelation as well as correlation between carriers on the same route. For robustness purposes, we ran specification (2) clustering by time period. This level of clustering accounts for any arbitrary correlation in the residuals by time period. Estimates of the coefficient on the output gap remain statistically significant at the 1 per cent level.

III(i). *Results*

Table II contains estimation results for both specifications, using the logarithm of the interquartile range and the Gini log-odds ratio as the dependent variable. We report results for all direct routes in the 17-year sample.[7]

[6] We measure price dispersion using the Gini log-odds ratio given by $G_{ij}^{lodd} = \ln\left(\frac{G_{ij}}{1-G_{ij}}\right)$, which produces an unbounded statistic. No results change when the log of the Gini coefficient is used instead. See Hayes and Ross [1998] for further discussion.

[7] This sample includes 154,407 carrier-route observations when using $ln(IQR)$ as the dependent variable and 156,038 carrier-route observations using the Gini log-odds ratio. The reason for the slight decrease in the number observations is that observations in which the interquartile range was equal to zero were necessarily dropped.

Table II
Full Sample Estimates

	$\ln(IQR)$		$Gini^{lodd}$	
YGAP	1.561***		1.145***	
	(0.199)		(0.104)	
-UR		2.271***		1.641***
		(0.286)		(0.152)
$\ln \widehat{HERF}$	0.261***	0.258***	0.079***	0.076***
	(0.043)	(0.043)	(0.023)	(0.023)
ln *FUEL*	0.046***	0.034**	–0.042***	–0.052***
	(0.015)	(0.016)	(0.009)	(0.009)
ln *COST*	0.135**	0.114**	0.450***	0.436***
	(0.061)	(0.057)	(0.030)	(0.030)
Observations	154407	153706	156038	155331

Notes: All regressions include carrier-route-specific dummies and quarter dummies. Standard errors are in parentheses and are clustered by route to account for both autocorrelation and correlation between carriers on the same route. One, two and three asterisks indicate significance at the 10 per cent, 5 per cent, or 1 per cent significance level, respectively.

The effect of a rise in the business cycle—as measured by the output gap—on price dispersion is positive and significant at the 1 per cent significance level. The estimate indicates that a one percentage point rise in the output gap (i.e., from 0.01 to 0.02) is associated with an increase in the interquartile range by 1.56 per cent and the Gini log-odds ratio by 0.011. The results from the second specification are similar to the first, indicating that a decrease in the unemployment rate is associated with an increase in the amount of price dispersion on a given route.[8] A one percentage point fall in the unemployment rate is associated with a 2.27 per cent increase in the interquartile range.

A look at the estimates from the percentile regressions in Table III sheds further light on the manner in which price dispersion follows the business cycle. The estimates show that the output gap is positively correlated with the 90th-percentile price level but is not positively correlated with the 10th-percentile price level. An increase in the output gap by one percentage point is associated with a 1.16 per cent increase in the 90th percentile price, but is not correlated with the 10th percentile price. Similarly, a fall in the unemployment rate by 1 percentage point is associated with a 1.37 per cent increase in the 90th percentile price, while there is a statistically significant, but small –0.339 per cent negative correlation between the unemployment rate and the 10th percentile price.

As in Gerardi and Shapiro [2009], we find that the effect of a decrease in competition—as measured by an increase in market concentration $\ln \widehat{HERF}$—on price dispersion is positive and significant at the 1 per cent

[8] This sample includes 153,706 carrier-route observations when using *ln*(*IQR*) as the dependent variable and 155,331 carrier-route observations using the Gini log-odds ratio. We have fewer observations in this specification because we do not have unemployment information for American Samoa or St. Thomas.

TABLE III
FULL SAMPLE ESTIMATES: PERCENTILES

	ln(90)		ln(10)	
YGAP	1.157***		–0.031	
	(0.104)		(0.061)	
-UR		1.373***		–0.339***
		(0.151)		(0.105)
$\ln \widehat{HERF}$	0.299***	0.296***	0.218***	0.218***
	(0.029)	(0.029)	(0.015)	(0.014)
ln *FUEL*	0.034***	0.024**	0.097***	0.097***
	(0.011)	(0.012)	(0.005)	(0.005)
ln *COST*	0.292***	0.276***	–0.015	–0.016
	(0.034)	(0.033)	(0.016)	(0.015)
Observations	156038	155331	156038	155331

Notes: All regressions include carrier-route-specific dummies and quarter dummies. Standard errors are in parentheses and are clustered by route to account for both autocorrelation and correlation between carriers on the same route. One, two, and three asterisks indicate significance at the 10 per cent, 5 per cent, or 1 per cent significance level, respectively.

significance level. There also appears to be interesting dynamics occurring on the cost side. Fuel costs seem to filter into both the 10th percentile prices and 90th percentile prices, while the slower moving operating costs filter only into the 90th percentile prices. There are many plausible stories that could explain this result. One possibility may be that carriers simply find it easier to pass costs on to the more price-insensitive consumers as they are more likely to lose the more price-sensitive consumers to competition.

As an additional exercise, we split our sample between legacy carriers and low-cost carriers (LCC's). Legacy carriers tend to implement different pricing strategies compared to the LCC's, so it is important to assess whether the type of carrier plays an important role in how price dispersion varies with the business cycle. For instance, some legacy carriers offer 'economy-plus,' which offers passengers more leg room, separate access through security, and/or early boarding. To determine whether these different types of carriers actually price differently over the business cycle, we re-estimate the main econometric specification for each sample separately. The estimates divided by carrier type are reported in the top panel of Table IV and show that most of the effects from the business cycle on price dispersion in the full sample of routes stem from the legacy carriers. The effect of the output gap on the interquartile range is slightly larger than two times the magnitude in the sample of legacy carriers ($\hat{\beta}_1 = 2.335$ compared to the estimated effect in the sample of LCC's $\hat{\beta}_1 = 0.965$).

Overall, the fixed-effects, panel estimates provide evidence of a positive relationship between the business cycle and price dispersion in the airline industry. Furthermore, the results show that the pro-cyclicality of price dispersion is largely driven by prices near the top of the price distribution.

TABLE IV
PANEL ESTIMATES BY CARRIER TYPE

	Legacy		LCC	
YGAP	2.335***		0.965***	
	(0.224)		(0.175)	
-UR		2.677***		1.747***
		(0.317)		(0.276)
ln $\widehat{HERF}$	0.191***	0.185***	0.311***	0.315***
	(0.048)	(0.048)	(0.039)	(0.039)
ln *FUEL*	–0.006	–0.027	0.225***	0.215***
	(0.016)	(0.016)	(0.013)	(0.012)
ln *COST*	0.069	0.029	0.579***	0.597***
	(0.045)	(0.045)	(0.042)	(0.043)
Observations	105636	104994	40941	40926

Notes: The dependent variable is the logarithm of the interquartile range. All regressions include carrier-route-specific dummies and quarter dummies. Standard errors are in parentheses and are clustered by route to account for both autocorrelation and correlation between carriers on the same route. One, two, and three asterisks indicate significance at the 10 per cent, 5 per cent, or 1 per cent significance level, respectively.

IV. DISCUSSION

In this section, we discuss some potential explanations for our empirical findings. Because the BTS does not provide detailed information on specific ticket or demographic characteristics, we cannot unequivocally single out any one specific pricing mechanism. However, we are able to provide empirical and theoretical evidence that favors certain explanations over others. We focus on two widely discussed theories of price dispersion in the airline industry: price discrimination and stochastic-demand pricing.

IV(i). *Second-Degree Price Discrimination*

The practice of price discrimination is one of the leading explanations for price dispersion in the airline industry. Airlines implement price discrimination techniques by segmenting heterogeneous groups of consumers and charging them distinct prices for a homogeneous product. Advance purchase requirements, non-refundable tickets, and Saturday-night layovers are a few examples of restrictions that airlines use to identify passengers with different price elasticities of demand. Since high-income or business consumers tend to place a high value on their time, they are more likely to purchase more expensive tickets without such restrictions. By making use of these techniques, airlines are able to separate price-sensitive travelers from price-insensitive travelers.

Using a parsimonious framework of second-degree price discrimination, we illustrate below that pro-cyclical price dispersion may be a side-effect of second-degree price discrimination. Specifically, under plausible assumptions of the utility function, a price discriminatory pricing policy implies that prices in the upper tail of the price distribution will be more sensitive to aggregate income fluctuations than prices in the lower tail. This suggests that price dispersion will positively covary with aggregate income when firms are price discriminating between consumers with different willingness-to-pay.

IV(i)(a). *Consumers.* Consider a simple model where consumers differ in their level of income, y. A consumer solves the following constrained utility maximization problem:

$$\max_{d\in\{0,1\}} d\cdot x+u(m) \tag{3}$$

subject to:

$$y=m+d\cdot p$$

where x represents the valuation of the ticket, which for now we assume is a constant. The variable m represents the numeraire commodity and d represents the consumer's decision to buy or not buy the good. Note that $u(\cdot)$ is the functional form representing the manner in which the consumer values the numeraire commodity relative to the discrete good, and we assume that it displays the conventional properties: $u'(y) > 0$ and $u''(y) < 0$. It follows that the indirect utility function for the case in which the consumer purchases the discrete good ($d = 1$) is given by:

$$U=x+u(y-p). \tag{4}$$

As in Tirole [1988], we make the assumption that a consumer's income is very large relative to the consumer's valuation, x, and subsequently to the equilibrium price charged. This allows us to take a first-order Taylor expansion around $p^* = 0$, which under the assumption that $y - p \approx y$, yields:

$$U=x+u(y)-u'(y)p. \tag{5}$$

It follows that for a given consumer to be better off consuming the good, it must be the case that $x + u(y) - u'(y)p \geq u(y)$, which means demand for the good is:

$$d(p)=\begin{cases}1 & if\ \ p\leq \dfrac{x}{u'(y)}\\[2ex] 0 & if\ \ p>\dfrac{x}{u'(y)}\end{cases} \tag{6}$$

IV(i)(b). *Firm Behavior.* To simplify the firm's problem, we assume two types of consumers and two types of tickets. The results below can easily be generalized to N types of consumers and N types of tickets. We assume there exists a share α of high income consumers with income y_h and a share $1-\alpha$ of low income consumers with income y_l. Quality is indexed by v, and we assume that there exists a high quality ticket, x_1, and a low quality ticket, x_2. For instance, $v = 2$ indicates a ticket that has an advance purchase requirement or Saturday-night stayover requirement, while $v = 1$ indicates a less restrictive ticket. It follows that with positive time costs, the net quality of $v = 1$ will be higher than that of $v = 2$ such that $x_1 > x_2$.

The firm's problem in the two-consumer-type case is to maximize profits given consumer demand derived above. The firm has the option to separate the market by offering different types of tickets. To obtain a separating equilibrium, the firm must be able to separate the market and also find it profit-maximizing to do so. It follows from Mussa and Rosen [1978] and Tirole [1988] that optimal incentive-compatible prices satisfy:

$$p_1^* = b_h x_1 - (b_h - b_l)x_2 \tag{7}$$

$$p_2^* = b_l x_2. \tag{8}$$

where $b_h = \frac{1}{u'(y)|_{y=yh}}$ and $b_l = \frac{1}{u'(y)|_{y=yl}}$. As the high-income consumer values x_2 more than the low-income consumer, the firm must lower the price of x_1 to dissuade the high-income consumer from deviating and purchasing the lower quality ticket, x_2. Specifically, the price is lowered by the extra utility the high-income consumer would have received over the low-income consumer by consuming x_2, $(b_h - b_l)x_2$. This lower price ensures that the high-income consumer does not purchase x_2 instead of x_1 (this ensures that the equilibrium is incentive compatible).[9]

Price sensitivities to a change in income, y, will be:

$$\frac{\partial p_1^*}{\partial y} = A_h b_h (x_1 - x_2) + A_l b_l x_2 \tag{9}$$

$$\frac{\partial p_2^*}{\partial y} = A_l b_l x_2 \tag{10}$$

[9] Maskin and Riley [1984] deal with a more general case where there is a choice over both quality and quantity. The authors show that the optimal quantity level is a function of the underlying quality. This implies that when consumers are allowed to make choices over both quantity and quality, it is optimal for the monopolist to offer a price schedule such that for a given quality level, he offers a unique quantity level associated with it, so that in fact consumption choices are made discrete. Hence *vis à vis* our discrete choice framework, the necessary correction would be taking into account the possibility of discrete quantity differences between the two consumer types.

where $A_h = \left[-\frac{u''(y)}{u'(y)}\right]_{y=y_h}$ and $A_l = \left[-\frac{u''(y)}{u'(y)}\right]_{y=y_l}$ are Arrow-Pratt measures of absolute risk-aversion (ARA) evaluated at income levels y_h and y_l, respectively. As the ARA will be positive as long as consumers have diminishing marginal utility of income, it follows from (7) and (8) that prices at the upper end of the distribution will be more sensitive to income shocks than prices in the bottom portion of the distribution. By contrast, if the firm chooses a uniform pricing strategy and wishes to sell to all consumers (i.e. both high and low income consumers), it must set a price in accordance with the low income consumer's preference parameter, b_l. For instance, if the firm chooses to sell only the low quality ticket, it would set a uniform price $b_l x_2$. If it chooses to sell only the high quality ticket it would set a uniform price $b_l x_1$. In the former case, average price sensitivity under a uniform pricing strategy will be less than the average price sensitivity under a discriminatory pricing strategy as long as $\alpha > 0$. In the latter case, this will be true only if $\alpha > \frac{A_l b_l}{A_h b_h}$.

The price range between the high and low price ticket (a measure of price dispersion similar to the interquartile range) is:

(11) $$D = p_1 - p_2 = b_h(x_1 - x_2).$$

It follows that the elasticity of price dispersion relative to a change in aggregate income, y, is:

(12) $$\varepsilon_{D,y} = \left[\frac{\partial D}{\partial y}\frac{y}{D}\right]_{y=y_h} = \left[-\frac{u''(y)}{u'(y)}y\right]_{y=y_h}$$

which is simply the coefficient of relative risk aversion (CRA) evaluated at y_h. As long as the CRA is positive (i.e., diminishing marginal utility of income), price dispersion will widen with an increase in aggregate income. It is important to note that this result is also sensitive to the choice of the utility function. Specifically, it will only hold for utility functions with the property that the consumer's willingness-to-pay increases jointly with both quality and income.

The model also shows that second-degree price discrimination may cause pro-cyclical price dispersion for reasons other than relative movements in price elasticities. For instance, equation (11) implies that price dispersion will follow the business cycle if there are relative movements in the non-price attributes of the good, $x_1 - x_2$, over the business cycle. This could happen if time costs are cyclical, or more generally speaking, if there are complementarities with certain ticket characteristics and business cycle conditions. Overall, this framework shows that price discriminatory tactics can cause price dispersion to widen during economic booms due to (1) movements in price elasticities or (2) movements in the non-price attributes of the good (that is, x). Since the DOT data does not include many of the ticket characteristics and because we do not have demographic information for the ticket purchasers, we cannot distinguish between these two effects.

IV(i)(c). *An Empirical Exercise.* To address the role of price-discriminatory behavior in generating pro-cyclical price dispersion, we perform an additional empirical exercise. Specifically, we assess the impact of consumer heterogeneity on the cyclicality of price dispersion by estimating equations (1) and (2) on two subsamples of routes: a sample of routes that are characterized by significant consumer heterogeneity in willingness to pay, as well as a sample of routes in which there is a more homogeneous consumer base. *Ceteris paribus*, there should be more opportunities to price discriminate in the former sample due to the presumed larger difference in willingness to pay of the consumers. To be clear, this is not a formal test of the theoretical model, but rather an additional specification of our empirical analysis of Section 3.1.

We decompose the full sample of routes into 'big-city' routes and 'leisure' routes, which we believe correspond to markets with heterogeneous consumer bases and markets with more homogeneous consumer bases, respectively. Since routes between large cities tend to attract both business and leisure travelers, they tend to have a bimodal distribution of prices while routes to largely leisure destinations, such as islands and beaches, tend to have unimodal price distributions and lower median prices.[10] Thus, airlines may have more opportunities to implement price discrimination strategies on these big-city routes since they include relatively more high income, business consumers. Furthermore, note that equation (12) implies that dispersion on big-city routes will be more sensitive to the cycle if the utility over income displays increasing relative risk aversion and less sensitive to the cycle if it displays decreasing relative risk aversion.[11] Table I shows summary statistics for the explanatory variables in each of these two samples. The table shows that costs are similar between the two samples, however, price dispersion is much larger in the sample of big-city routes.

Table V contains estimates of the correlation between the interquartile range of the price distribution for a carrier-route observation and the business cycle for big-city routes versus leisure routes. The estimates show that price dispersion is more closely tied with the output gap and the average city-wide unemployment rate for the big-city route sample than the leisure sample. For instance, in the big-city route sample, a one percentage point rise in the output gap is associated with a 2.7 per cent increase in the interquartile range, while it is associated with a statistically insignificant

[10] For a full list of the cities in each sample as well as a detailed description of how these subsamples are created, see Gerardi and Shapiro [2009].

[11] This can be seen by taking the derivative of (12) with respect to y_h.

TABLE V
PANEL ESTIMATES BY ROUTE TYPE

	Big-City		Leisure	
YGAP	2.732***		0.194	
	(0.352)		(0.467)	
-UR		2.756***		2.248***
		(0.482)		(0.605)
$\ln \widehat{HERF}$	0.334***	0.329***	0.195***	0.209***
	(0.078)	(0.079)	(0.048)	(0.053)
ln *FUEL*	–0.054**	–0.080***	0.190***	0.194***
	(0.024)	(0.026)	(0.027)	(0.031)
ln *COST*	0.093	0.047	–0.085	–0.075
	(0.070)	(0.070)	(0.162)	(0.151)
Observations	43614	43614	35312	34611

Notes: The dependent variable is the logarithm of the interquartile range. All regressions include carrier-route-specific dummies and quarter dummies. Standard errors are in parentheses and are clustered by route to account for both autocorrelation and correlation between carriers on the same route. One, two, and three asterisks indicate significance at the 10 per cent, 5 per cent, or 1 per cent significance level, respectively.

0.19 per cent increase in the leisure route sample. The effect of the unemployment rate on price dispersion is slightly larger in the big-city sample, a coefficient of 2.8 compared to 2.3 in the leisure sample. This smaller difference in magnitude of the coefficients may be due to the fact that, as opposed to the output gap, the unemployment rate is specific to the economic conditions at the endpoint cities.

IV(ii). *Stochastic-Demand Pricing*

Another important theory regarding the existence of price dispersion is that of stochastic-demand pricing. If the carrier is constrained by capacity, then as more flights reach full capacity, the expense of an additional passenger becomes very large as either a bigger aircraft or an extra flight is needed to supply the extra seat-mile. Eden [1990] shows that effect can induce price dispersion to rise in periods of peak demand when full capacity is reached.

In discussing the effect of capacity constraints on pricing, it is useful to decompose marginal cost into its two primary components, which we refer to as the passenger cost and the capacity cost. If the aircraft is not operating at full capacity, then marginal cost is simply equal to the passenger cost; the cost of adding an additional passenger to the airplane. This cost is mostly made up of the extra fuel required to transport the additional weight of the passenger, while other, lesser components include the in-flight costs of serving the additional passenger (i.e. meals, snacks, etc.). However, if the airplane is operating at full capacity, then marginal cost is equal to the direct cost of an additional passenger as well as the more substantial cost of an additional flight. This cost is incurred regardless of whether or not seats on the airplane are filled with passengers, while the passenger cost is only incurred on seats that are sold. This implies that marginal cost at the route level is given by,

$$c_{ij} = \begin{cases} \beta_{ij}, & \text{if capacity is not reached} \\ \beta_{ij} + \lambda_{ij}, & \text{if capacity is reached} \end{cases}$$

where β_{ij} is the cost of serving an additional passenger one mile on route j by carrier i, and λ_{ij} is the cost of an additional flight (in seat-miles).

If airlines account for stochastic demand concerns in their pricing decisions, then aggregate demand fluctuations could alter a firm's expected probability of selling a ticket, and subsequently alter the 'effective' capacity cost. In particular, if *ex-ante* the carrier is uncertain about the level of demand for a flight, then under price-setting commitments and costly capacity, profit-maximizing behavior induces a distribution of prices rather than a single price. The intuition is that if the firm were allowed to change price after the realization of the state, then it would set a low price in the low-demand state and a high price in the high-demand state. However, because the firm must commit to a menu of prices *ex-ante*, its profit maximizing strategy is to assign multiple prices to specified quantities of the good. That is, if a firm must pay costs irrespective of whether or not its output is sold, then it has a large incentive to set higher prices on goods that are less likely to be sold.

Eden [1990] formalized a model in a setting of perfect competition where there is uncertainty regarding the number of agents who will show up to exchange goods in the marketplace. In such a setting, goods are characterized by the probability that they will be sold, and in equilibrium, firms face a tradeoff between price and the probability of sale. In the model, equilibrium prices are given by the condition,

$$p_s = \beta + \underbrace{\frac{\lambda}{prob(sale)_s}}_{\lambda_s^{eff}} \tag{13}$$

where p_s is the price of the *sth* good, β is an operating cost that the firm must pay for each good that it sells, λ is the unit capacity cost, and $prob(sale)$ is the probability that good s is sold. The second term on the right-hand side of the equation can be interpreted as an 'effective' capacity cost of good s, λ_s^{eff}. This term implies that in competitive equilibrium, firms are indifferent between selling a high-priced good with low probability and selling a low-priced good with high probability. Dana [1999] extended Eden's model to monopoly and oligopoly market structures. With stochastic demand, the monopolist sets a higher price for a good that sells only in high demand states since its effective cost is higher.

In this setting, when the carrier commits to prices *ex-ante*, the highest priced tickets—tickets with the highest effective capacity cost—are not purchased until demand rises sufficiently high to purchase all of the low

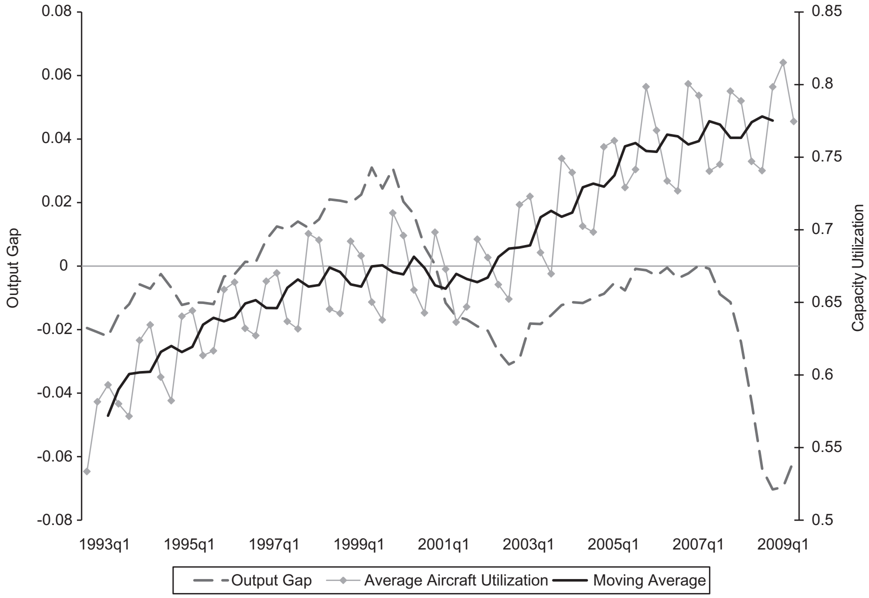

Figure 4
Aircraft Utilization Rate

Notes: Depicted as a grey line is aircraft utilization, measured as total passengers divided by total available seats from the BTS T100 database, for each quarter in the sample period. The black line depicts a moving average of this measure. The output gap, as measured by the Congressional Budget Office (CBO) is depicted as a dashed line.

priced tickets. Thus, if the carrier is pricing solely with stochastic demand concerns, then peaks in aggregate demand will induce higher price dispersion through the higher effective capacity cost of the remaining seats on crowded aircrafts.

IV(ii)(a). *An Empirical Exercise.* To assess the empirical importance of stochastic-demand pricing in generating pro-cyclical price dispersion, we exploit the expected relationship between capacity utilization and price dispersion that would arise if stochastic demand played an important role in airline pricing tactics. Specifically, under stochastic-demand pricing, utilization should positively co-vary with price dispersion because high-priced tickets would be purchased only when aircrafts are near full capacity.

Figure 4 shows the mean aircraft capacity utilization rate over the sample period. Interestingly, utilization steadily increased over the course of the sample period, fluctuating with some seasonal variation. As a formal test, we control for the effects of stochastic-demand pricing on price dispersion by including a measure of carrier *i*'s utilization rate on route *j* in our estimation routine, $UTIL_{ijt}$. As this variable is potentially endogenous, we also include specifications using instrumental variables. Specifically, we instrument for $UTIL_{ijt}$ using the utilization of route j in period t, $UTIL_{ijt}$. This variable should be correlated with airline i's specific utilization rate due to variations in aggregate demand for route j. As Figure 4 makes apparent, it may be important to remove low frequency components from the utilization variable. Thus, we also include specifications with a de-trended measure of aircraft utilization, $UTIL_{dt}$.[12]

We report results of this exercise using two measures of price dispersion. Estimates using the logarithm of the interquartile range are reported in the top panel of Table VI and estimates with the Gini-log odds ratio are reported in the bottom panel. In all specifications, the coefficient on our measure of the business cycle is positive and statistically significant at the one per cent level. Thus, holding fixed aircraft utilization, price dispersion remains pro-cyclical. These estimates suggest that the pro-cyclicality of variation in airline price dispersion is likely not tied to variation in capacity utilization. In turn, this suggests that stochastic-demand pricing strategies do not explain our findings. While this analysis favors price discrimination as the explanation for the pro-cyclical nature of airline price dispersion, it is important to stress that we can only favor certain explanations over others due to certain limitations of the data. For instance, we use a monthly measure of capacity utilization at the carrier-route level, whereas ideally, we would like capacity utilization measured at the flight level. It is comforting to note, however, that our results correspond with recent work of Puller, Sengupta and Wiggins [2009] who use more granular ticket information.

V. CONCLUSION

In this paper we have documented that price dispersion is significantly pro-cyclical in the airline industry. We show that the empirical results are consistent with a parsimonious discrete-choice model of second degree price discrimination. In addition, we implement a few empirical exercises that provide support for this interpretation over others such as stochastic demand pricing or pro-cyclical variation in airline costs. With the available data, we cannot completely rule out other mechanisms that could create pro-cyclical behavior in price dispersion. One such mechanism is changes in consumer behavior over the business cycle. Even in the absence of variation in airlines' pricing strategies over the business cycle, if consumers simply purchase more expensive airline tickets with fewer restrictions in boom periods and less expensive tickets with more restrictions in more depressed

[12] The measure $UTIL_{dt}$ is measured by collecting the residuals from running a regression of utilization, *util*, on a cubic-polynomial time trend and quarter dummies.

TABLE VI
ESTIMATES WITH CAPACITY UTILIZATION CONTROL

ln (IQR)								
	OLS				IV			
YGAP	1.896***	2.161***			1.923***	2.196***		
	(0.213)	(0.217)			(0.216)	(0.227)		
-UR			2.687***	3.007***			2.721***	3.055***
			(0.281)	(0.287)			(0.279)	(0.275)
$UTIL$	−1.122***		−1.121***		−1.211***		−1.213***	
	(0.084)		(0.081)		(0.101)		(0.096)	
$UTIL_{dt}$		−1.217***		−1.208***		−1.286***		−1.287***
		(0.077)		(0.073)		(0.109)		(0.104)
ln $\widehat{HERF}$	0.239***	0.236***	0.235***	0.231***	0.241***	0.240***	0.237***	0.234***
	(0.040)	(0.040)	(0.040)	(0.040)	(0.040)	(0.040)	(0.040)	(0.040)
ln $FUEL$	0.120***	0.055***	0.104***	0.037**	0.126***	0.055***	0.110***	0.037***
	(0.012)	(0.013)	(0.013)	(0.015)	(0.011)	(0.013)	(0.012)	(0.014)
ln $COST$	0.173***	0.096*	0.148***	0.067	0.176***	0.093*	0.150***	0.063
	(0.061)	(0.055)	(0.057)	(0.052)	(0.061)	(0.055)	(0.057)	(0.051)
Observations	154390	154390	153689	153689	154390	154390	153689	153689
$Gini^{lodd}$								
	OLS				IV			
YGAP	1.235***	1.375***			1.233***	1.301***		
	(0.106)	(0.110)			(0.105)	(0.108)		
-UR			1.752***	1.923***			1.751***	1.833***
			(0.155)	(0.160)			(0.155)	(0.158)
$UTIL$	−0.305***		−0.303***		−0.299***		−0.298***	
	(0.036)		(0.036)		(0.043)		(0.043)	
$UTIL_{dt}$		−0.475***		−0.469***		−0.321***		−0.320***
		(0.040)		(0.040)		(0.047)		(0.047)
ln $\widehat{HERF}$	0.072***	0.068***	0.069***	0.065***	0.072***	0.071***	0.069***	0.068***
	(0.023)	(0.022)	(0.023)	(0.022)	(0.023)	(0.022)	(0.023)	(0.023)
ln $FUEL$	−0.022***	−0.039***	−0.033***	−0.051***	−0.023***	−0.040***	−0.033***	−0.051***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)	(0.008)	(0.008)	(0.008)
ln $COST$	0.460***	0.434***	0.445***	0.417***	0.460***	0.439***	0.445***	0.423***
	(0.030)	(0.029)	(0.030)	(0.028)	(0.031)	(0.029)	(0.030)	(0.028)
Observations	156021	156021	155314	155314	156021	156021	155314	155314

Notes: All regressions include carrier-route-specific dummies and quarter dummies. IV estimates indicate that $UTIL$ and $UTIL_{dt}$ were instrumented using utilization at the route-quarter level. Standard errors are in parentheses and are clustered by route to account for both autocorrelation and correlation between carriers on the same route. One, two, and three asterisks indicate significance at the 10 per cent, 5 per cent, or 1 per cent significance level, respectively.

periods, then our measured degree of price dispersion would be expected to co-vary with the business cycle. Future research on this topic will hopefully employ more detailed data that will have the ability to distinguish between causal mechanisms, and thus determine if price discrimination is the most important driver of price dispersion at the business cycle frequency. Another potentially fruitful avenue for future research is to analyze price dispersion in other industries. If pro-cyclical price dispersion is largely due to price discrimination tactics, then we would expect to find pro-cyclical price dispersion in industries in which firms rely on price discriminatory strategies, such as hotels, stadiums, restaurants, theaters, yellow-page advertising, cement, and personal computers.

Another interesting extension of this study would be to assess whether price discriminatory tactics act to accentuate the degree to which airline profits fluctuate over the business cycle. Given the high volatility of profits over the course of the last two decades as well as the large number of bankruptcies by legacy carriers, the airline industry seems particularly sensitive to aggregate demand conditions. But while legacy carriers have struggled, the LCC's have somehow managed to stay profitable during this era. One possibility is that the large profit swings of legacy carriers relative to LCC's are, in part, attributable to differences in the reliance on price discriminatory tactics, as LCC's such as Southwest and JetBlue do not price discriminate to the same extent as legacy carriers.

APPENDIX A
VARIABLE DEFINITIONS

- $\ln P(k)_{ijt}$—The logarithm of the kth price percentile of carrier i on route j in period t, obtained from the DB1B.
- $\ln IQR_{ijt}$—The logarithm of the interquartile range, given by $P(75)_{ijt} - P(25)_{ijt}$, where $P(k)_{ijt}$ is the price percentile of carrier i on route j in period t, obtained from the DB1B.
- $Gini_{ijt}^{lodd}$—The Gini log-odds ratio, given by $G_{ijt}^{lodd} = \ln\left(\frac{G_{ij}}{1-G_{ij}}\right)$, where G_{ijt} is the Gini coefficient of carrier i's price distribution on route j in period t, calculated using data from DB1B.
- $\ln HERF_{jt}$—The logarithm of the Herfindahl-Hirschman index of route j in period t, calculated using passenger shares obtained from the DB1B.
- $YGAP_t$—The log of nominal GDP in period t minus the log of nominal potential GDP in period t, as measured by the Congressional Budget Office (CBO).
- UR_{jt}—The average metropolitan unemployment rate in period t of the origin and destination state of route j, obtained from Bureau of Labor Statitics (BLS).
- $\ln FUEL_{it}$—The average cost per gallon fuel by carrier i in period t, obtained from the BTS P-52 database.
- $\ln COST_{it}$—Total operating costs minus total fuel costs divided by total seat-miles for carrier i in period t, obtained from the BTS P-52 database.

- $UTIL_{ijt}$—The capacity utilization rate of carrier i on route j in period t measured by total passengers divided by total seats. Obtained from the T-100 database.
- $UTIL_{dt}$—The de-trended capacity utilization rate of carrier i on route j in period t measured as the residual from the regression of $util_{ijt}$ on a cubic-polynomial time trend and quarter dummies.

Instruments

- $\ln PASSRTE_{jt}$—The logarithm of total enplaned passengers on route j in period t from the T-100 Domestic Segment Databank.
- *IRUTHERF*—This instrument is identical to one used by Borenstein and Rose [1994]. This variable is the square of the fitted value for $MKTSHARE_{ijt}$ from its first-stage regression, plus the rescaled sum of the squares of all other carrier's shares. See Borenstein and Rose [1994] for a more detailed explanation. It is equal to $\widehat{MKTSHARE}_{ijt}^2 + \dfrac{HERF_{jt} - MKTSHARE_{ijt}^2}{(1 - MKTSHARE_{ijt})^2} * (1 - \widehat{MKTSHARE}_{ijt})^2$.
- *GENSP*—$\dfrac{\sqrt{ENP_{j1} * ENP_{j2}}}{\sum_k \sqrt{ENP_{k1} * ENP_{k2}}}$, where k indexes all airlines, j is the observed airline, and ENP_{k1} and ENP_{k2} are airline k's average quarterly enplanements at the two endpoint airports. This instrument is similar to one used by Borenstein and Rose [1994], with the difference being that Borenstein and Rose use average daily enplanements, while we use average quarterly enplanements, as a result of data availability. Data on enplanements were obtained from the T-100 Domestic Segment Databank.

APPENDIX B
DATA CONSTRUCTION

In this appendix, we discuss our methods and assumptions involved in constructing our panel of airline-route ticket observations from the DB1B and T-100 Domestic Segment databases maintained by the BTS at their online website, Transtats. There are three different sub-components to the DB1B data set. They are market data, coupon data and ticket data, and we combine variables from all three sources.[13]

We use only domestic, coach-class itineraries and keep only tickets containing direct flights.[14] Direct flights typically account for 30 per cent of the itineraries in the DB1B over the course of our sample, with no apparent trend.

The BTS includes a variable that describes the reliability of each ticket price ('dollar cred'). The variable takes on a value of 0 if the fare is of questionable magnitude, based on a set of limits defined by the BTS, and it takes a value of 1 if it is credible. We drop all tickets for which this variable takes a value of zero.

The DB1B also provides limited information regarding the fare class of each ticket. Each ticket is labeled as either coach-class, business-class or first-class, and we elimi-

[13] For further reference, see the BTS's website *http://www.transtats.bts.gov*.

[14] The sample of direct flights encompasses both non-stop flights and flights in which there is a stop but no change of plane.

nated all first-class and business-class itineraries. Unfortunately, the DB1B does not have any direct way of identifying frequent-flyer tickets, but there are indirect methods that have been used in the previous literature, and we follow these in our analysis. First, we drop all fares coded as 0. Next, we dropped all fares that are less than or equal to $20 ($10 for one-way tickets).

In addition to eliminating frequent-flyer tickets and higher-class tickets, we also eliminate tickets in which the operating and ticketing carriers are different due to code sharing arrangements. Code sharing is a practice where a flight operated by an airline is jointly marketed as a flight for one or more other airlines. Due to the uncertainty regarding the actual airline who is setting the price schedule in such an arrangement, we decided to eliminate these itineraries. Code sharing first appears in the data in 1998:Q1. On average, approximately 80 per cent of the original number of direct tickets in the DB1B is retained in the analysis.

After filtering the ticket data for each quarter of the DB1B, we combined tickets from all 55 quarters and collapsed the data into airline-route observations. For example, if we had 10,000 United Airline tickets between Boston and Los Angeles in 1993:Q1, we calculated summary statistics (such as the Gini coefficient), and collapsed the data into a single observation corresponding to a United Airlines flight between Boston and Los Angeles in 1993:Q1.

The merge between the DB1B and T-100 Segment databases was not exact (around 45 per cent matched). First, since the DB1B does not provide complete coverage for all airlines and routes, there are a number of direct routes in the T-100 data that we do not find in the DB1B (especially low-volume routes). Second, the DB1B does not allow us to distinguish between a non-stop, direct ticket and a ticket that involves a stop without a plane change. For example, if a passenger takes a flight from Boston to Orlando that stops in Atlanta, but does not involve a plane change, his itinerary will look identical to that of a passenger who flies from Boston to Orlando without any stops. For this reason, we identified some airline routes as direct in the DB1B, that are not non-stop, and therefore do not have segment information in the T-100 data. While we lose many airline-route observations during the merge as a result, we believe that this merge actually provides a nice filter, since we would ideally like to use only non-stop, direct flights. Thus, by merging data between the DB1B and the T-100, we likely eliminate a large proportion of flights that are direct, but not non-stop due to a plane change.

In an effort to eliminate possible coding errors, we drop certain airline-route observations from the data that we believe do not have adequate coverage to calculate reliable price dispersion statistics. We drop any airline-route observation that does not have at least 100 passengers in the DB1B. Furthermore, for each airline route observation, we calculate the average number of passengers over time in both the DB1B and the T-100 Segment databases. If the number of passengers on an airline route in a given quarter falls below 25 per cent of its mean over time in one of the databases, but not in the other, then we drop the observation from our data, on the basis that its value is most likely measurement error. However, if the number of passengers on an airline route in a given quarter falls below 25 per cent of its mean in both the DB1B and the T-100 Segment databases, then we keep the observation in our data.

Finally, we addressed the issue of 'double counting.' Since we defined a route as a directional trip in our data, any round-trip ticket would count twice. For example, a round-trip fare from Boston to San Francisco would appear twice in the data—once as BOS-SFO and once as SFO-BOS. Since this would have no effect on the consistency of our estimates, but a significant effect on the size of our standard errors, we chose to drop one of the directions. Of course, the drawback of this assumption is that some one-way fares were dropped from the data as a result. In our judgment, the first issue outweighed the second issue.

REFERENCES

Aguirre, I.; Cowan, S. and Vickers, J., 2010, 'Monopoly Price Discrimination and Demand Curvature,' *American Economic Review*, 100(4) pp. 1601–1615(15).

Aizcorbe, A. and Shapiro, A.H., 2010, 'Implications of Consumer Heterogeneity on Price Measures for Technology Goods.' BEA working paper, Because of Economic Analysis, Department of Commerce, Washingto, D.C., U.S.A.

Baye, M.; Morgan, J. and Scholten, P., 2004, 'Price Dispersion in the Small and in the Large: Evidence from an Internet Price Comparison Site,' *The Journal of Industrial Economics*, 52(4), pp. 463–496.

Borenstein, S. and Rose, N., 1994, 'Competition and Price Dispersion in the U.S. Airline Industry,' *The Journal of Political Economy*, 102(4), pp. 653–683.

Bils, M., 1989, 'Pricing in a Customer Market,' *Quarterly Journal of Economics*, 104(4), pp. 699–718.

Burdett, K. and Judd, K., 1983, 'Equilibrium Price Dispersion,' *Econometrica*, 51(4), pp. 955–969.

Busse, M. and Rysman, M., 2005, 'Competition and Price Discrimination in Yellow Pages Advertising,' *The RAND Journal of Economics*, 36(2), pp. 378–390.

Caves, D.; Christensen, L. R. and Tretheway, M. W., 1984, 'Economies of Density versus Economies of Scale: Why Trunk and Local Service Airline Costs Differ,' *Rand Journal of Economics*, 15(4), pp. 471–489.

Chevalier, J. and Scharfstein, D., 1996, 'Capital-Market Imperfections and Countercyclical Markups: Theory and Evidence,' *American Economic Review*, 86(4), pp. 703–725.

Chevalier, J.; Kashyap, A. and Rossi, P., 2003, 'Why Don't Prices Rise During Periods of Peak Demand? Evidence from Scanner Data,' *American Economic Review*, 93(1), pp. 15–37.

Dana, J., 1999, 'Equilibrium Price Dispersion Under Demand Uncertainty: The Roles of Costly Capacity and Market Structure,' *The RAND Journal of Economics*, 30(4), pp. 632–660.

Eden, B., 1990, 'Marginal Cost Pricing When Spot Markets Are Complete,' *The Journal of Political Economy*, 98(6), pp. 1293–1306.

Gaggero, A. and Piga, C., 2011, 'Airline Market Power and Intertemporal Price Dispersion,' Forthcoming *Journal of Industrial Economics*.

Gerardi, K. and Shapiro, A.H., 2009, 'Does Competition Reduce Price Dispersion? New Evidence from the Airline Industry,' *Journal of Political Economy*, 117(1), pp. 1–37.

Gillen, D.W.; Oum, T.H. and Tretheway, M.W., 1990, 'Airline Cost Structure and Policy Implications: A Multi-Product Approach for Canadian Airlines,' *Journal of Transport Economics and Policy*, 24(1), pp. 9–34.

Goldberg, P. and Verboven, F., 2001, 'The Evolution of Price Dispersion in the European Car Market,' *Review of Economic Studies*, 68, pp. 811–848.

Goolsbee, A. and Syverson,C., 2008. 'How Do Incumbents Respond to the Threat of Entry? Evidence from the Major Airlines,' *The Quarterly Journal of Economics*, 123(4), pp. 1611–1633.

Greenwald, B.; Stiglitz, J. and Weiss, A., 1984, 'Informational Imperfections in the Capital Market and Macroeconomic Fluctuations,' *American Economic Review Papers and Proceedings*, 74(2), pp. 194–99.

Hayes, K. and Ross, L., 1998, 'Is Airline Price Dispersion the Result of Careful Planning or Competitive Forces?' *Review of Industrial Organization*, 13, pp. 523–542.

Hendel, I. and Nevo., A., 2011, 'Intertemporal Price Discrimination in Storable Goods Markets,' NBER Working Paper No. 16988. National Bureau of Economic Research, Cambridge, Massachusetts, U.S.A.

Holmes, T., 1989, 'The Effect of Third Degree Price Discrimination in Oligopoly,' *American Economic Review*, 79(1), pp. 244–250.

Ito, H. and Lee, D., 2003, 'Low Cost Carrier Growth in the U.S. Airline Industry: Past, Present and Future,' Brown University Department of Economics Paper No. 2003-12; Brwn University, Providence, Rhode Island, U.S.A.

Klemperer, P., 1995, 'Competition when Consumers Have Switching Costs: An Overview with Applications to Industrial Organization, Macroeconomics and International Trade,' *Review of Economic Studies*, 62(4), pp. 515–539.

Leslie, P., 2004, 'Price Discrimination in Broadway Theater,' *The RAND Journal of Economics*, 35(3), pp. 520–541.

Maskin, E. and Riley, J., 1984, 'Monopoly with Incomplete Information,' *RAND Journal of Economics*, 15, pp. 171–96.

Miller, N. and Osborne, M., 2010, 'Competition among Spatially Differentiated Firms: An Empirical Model with an Application to Cement,' EAG Dissussion Paper 10-2, Economic Analysis Group, Antitrust Division, U.S. Department of Justice, Washington, D.C., U.S.A.

MacDonald, J., 2000, 'Demand, Information and Competition: Why Do Food Prices Fall at Seasonal Demand Peaks?' *Journal of Industrial Economics*, 48(1), pp. 27–45.

Mussa, M. and Rosen, S., 1978, 'Monopoly and Product Quality,' *Journal of Economic Theory*, 18, pp. 301–317.

Nevo, A. and Hatzitaskos, K., 2006, 'Why Does the Average Price Paid Fall During High Demand Periods?' NBER Working Paper No. 11572, National Bureau of Economic Research, Cambridge, Massachusetts, U.S.A.

Orlov, E., 2011, 'How Does the Internet Influence Price Dipsersion? Evidence from the Airilne Industry,' *Journal of Industrial Economics*, 59(1), pp. 21–37.

Puller, S.; Sengupta, A. and Wiggins, S., 2009, 'Testing Theories of Scarcity Pricing in the Airline Industry,' NBER Working Paper 15555, National Bureau of Economic Research, Cambridge, Massachusetts, U.S.A.

Rotemberg, J. and Saloner, G., 1986, 'A Super-game Theoretic Model of Business Cycles and Price Wars During Booms,' *American Economic Review*, 76(3), pp. 390–407.

Salop, S. and Stiglitz, J., 1982, 'The Theory of Sales: A Simple Model of Equilibrium Price Dispersion with Identical Agents,' *American Economic Review*, 72(5), pp. 1121–1130.

Shepard, A., 1991, 'Price Discrimination and Retail Configuration,' *The Journal of Political Economy*, 99(1), pp. 30–53.

Sorensen, A., 2000, 'Equilibrium Price Dispersion in Retail Markets for Prescription Drugs,' *Journal of Political Economy*, 108(4), pp. 833–850.

Stavins, J., 2001, 'Price Discrimination in the Airline Market: The Effect of Market Concentration,' *Review of Economics and Statistics*, 83(1), pp. 200–202.

Stiglitz, J., 1984, 'Price Rigidities and Market Structure,' *American Economic Review*, 74(2), pp. 350–355.

Tirole, J., 1988, *The Theory of Industrial Organization*, MIT Press, Cambridge, Massachasctts, U.S.A.

Warner, E. and Barsky, R., 1995, 'The Timing and Magnitude of Retail Store Markdowns: Evidence from Weekends and Holidays,' *Quarterly Journal of Economics*, 110(2), pp. 321–352.