

Forthcoming, *System Dynamics Review*

Cyclical Dynamics of Airline Industry Earnings

Kawika Pierson*

Atkinson Graduate School of Management, Willamette University, Salem OR

John Sterman

MIT Sloan School of Management, Cambridge, MA

ABSTRACT

Aggregate airline industry earnings have exhibited large amplitude cyclical behavior since deregulation in 1978. To explore the causes of these cycles we develop a behavioral dynamic model of the airline industry with endogenous capacity expansion, demand, pricing, and other feedbacks; and model several strategies industry actors have employed in efforts to mitigate the cycle. We estimate model parameters by maximum likelihood methods during both partial model tests and full model estimation using Markov chain Monte Carlo methods to establish confidence intervals. Contrary to prior work we find that the delay in aircraft acquisition (the supply line of capacity on order) is not a very influential determinant of the profit cycle. Instead we find that aggressive use of yield management—varying prices to ensure high load factors (capacity utilization)—may have the unintended effect of increasing earnings variance by increasing the sensitivity of profit to changes in demand.

KEYWORDS: Earnings cycles, Profit cycles, Airlines, Operational leverage, Capacity supply line, Yield management

* Corresponding author: kpierson@willamette.edu

Introduction

Researchers in system dynamics have studied cyclicity in industries and the economy for decades (Forrester, 1961; Meadows, 1970; Mass, 1975), and have generally concluded that profit cycles are caused by a failure to fully account for delays in the negative feedbacks controlling inventory, capacity acquisition, or other resources. Unfortunately, the low salience of capacity on order (Sterman, 1989; Sterman, 2000) together with long capacity lifetimes and high fixed costs often limit the implementation of strategies to mitigate the cycle, because managers can be reluctant to accept that such important decisions could have been detrimental (Ghaffarzadegan and Tajrishi, 2010; Goncalves, 2003).

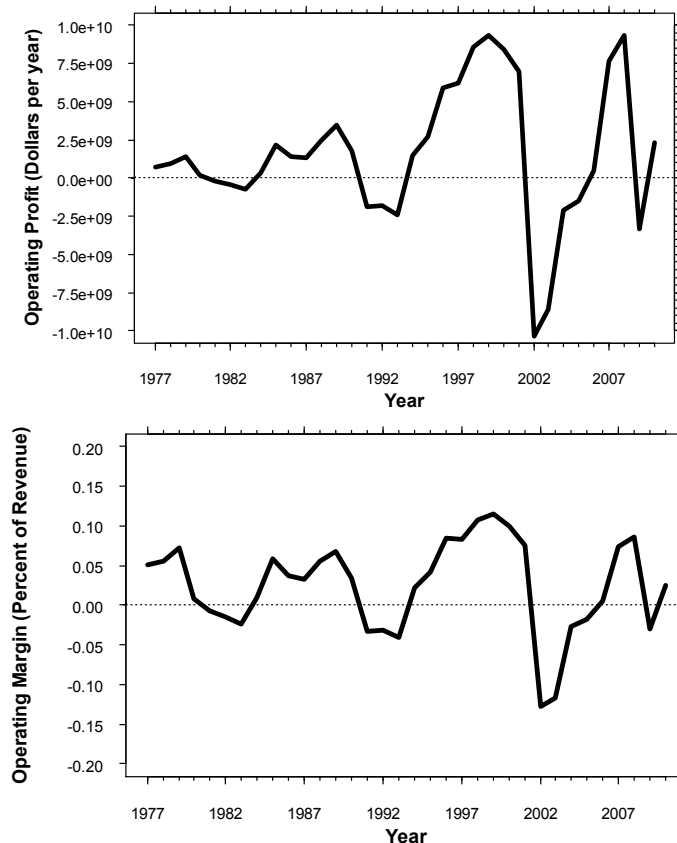


Figure 1: US airline industry operating profit and operating margin (profit/revenue)

Since deregulation in 1978 aggregate earnings of the US airline industry have fluctuated with an average peak-to-peak period of approximately ten years and a long run mean very close to zero (Hansman and Jiang, 2005), as shown in Figure 1. The amplitude of the cycle in profit margin (operating profit/revenue, a scale-free measure of profit fluctuations), has not diminished in the 35 years since deregulation. In this paper we build a model of the airline industry that examines the origin of the cycle. Airline industry cyclicity has been

addressed in the system dynamics literature (Liehr et al., 2001; Lyneis, 2000), but we expand the boundary of these models¹ to include an endogenous account of feedbacks omitted from some earlier work, including price setting, wages, and air travel demand. Including these feedbacks allows us to more closely represent the structure of the industry so as to better test policies designed to moderate the cycle. The model also includes structures representing yield management, mothballing, and ancillary revenues (e.g., baggage check fees) to address how existing strategic decisions influence profits and profit variability.

The airline industry is an excellent setting for research on profit cycles. The government requires airlines to report detailed information about their operations, and makes these data available to the public. By avoiding proprietary sources of data, we provide a fully documented model that scholars and industry professionals can use to better understand the dynamics of earnings cycles in general. We estimate model parameters via maximum likelihood methods, using both partial model tests (Homer, 2012) and full model estimation, and show how standard errors can be estimated efficiently in multivariate system dynamics models using Markov chain Monte Carlo methods.

Airlines are also advantageous as a research setting because of their importance. The Federal Aviation Administration (2011) estimates that commercial aviation contributes \$1.2 to \$1.3 trillion per year to the economy and generates between 9.7 and 10.5 million jobs in the US. Yet despite the importance of the industry, consistent profitability has been elusive. Industry analysts and experts are not blind to this pattern of behavior. Like their peers in other cyclical industries, they consistently argue either that specific events were the cause of the cycle turning points (e.g., recessions or terrorist attacks) or that new strategies will dampen the cycle in the future (Doganis, 2002). These arguments persist in the face of a history of strategies, such as mergers, leasing, yield management, and mothballing that have failed to stabilize aggregate profits.²

Consistent with prior system dynamics work, we find that the cycle arises from delays in negative feedbacks involving the mutual regulation of demand, capacity, load factor (capacity utilization), and prices. Unlike prior work, we find evidence that the cycle in capacity is strongly moderated by airline pricing policies, specifically the use of yield management, which increases the responsiveness of prices to variations in demand relative to capacity and boosts average load factors. However, sensitivity tests varying the strength of the yield management feedback suggest that in the aggregate, airline pricing decision rules increase operational leverage and the variance in profitability, and may place the industry near a global minimum of the risk-return space.

¹ Lyneis (2000) has a similar model boundary but is proprietary.

² Yield, the industry term for dollars per revenue passenger mile, and price are used interchangeably in this paper. Yield management is the process of “finding the optimal tradeoff between average price paid and capacity utilization” (Weatherford and Bodily, 1992).

Model simulations, together with the low average price to earnings ratio of airline stocks and the high incidence and cost of airline bankruptcies, suggest that airlines could potentially improve long-run shareholder value by adopting policies that pursue less vigorous yield management. The feasibility and full impacts of such policies for individual airlines may depend on competitive dynamics beyond the level of aggregation of the model however, so we close by discussing the limitations of our analysis and suggestions for future research to build on the results here.

Model Structure

Figure 2 shows a high level causal diagram summarizing the principal feedbacks captured by and the exogenous influences to the model.

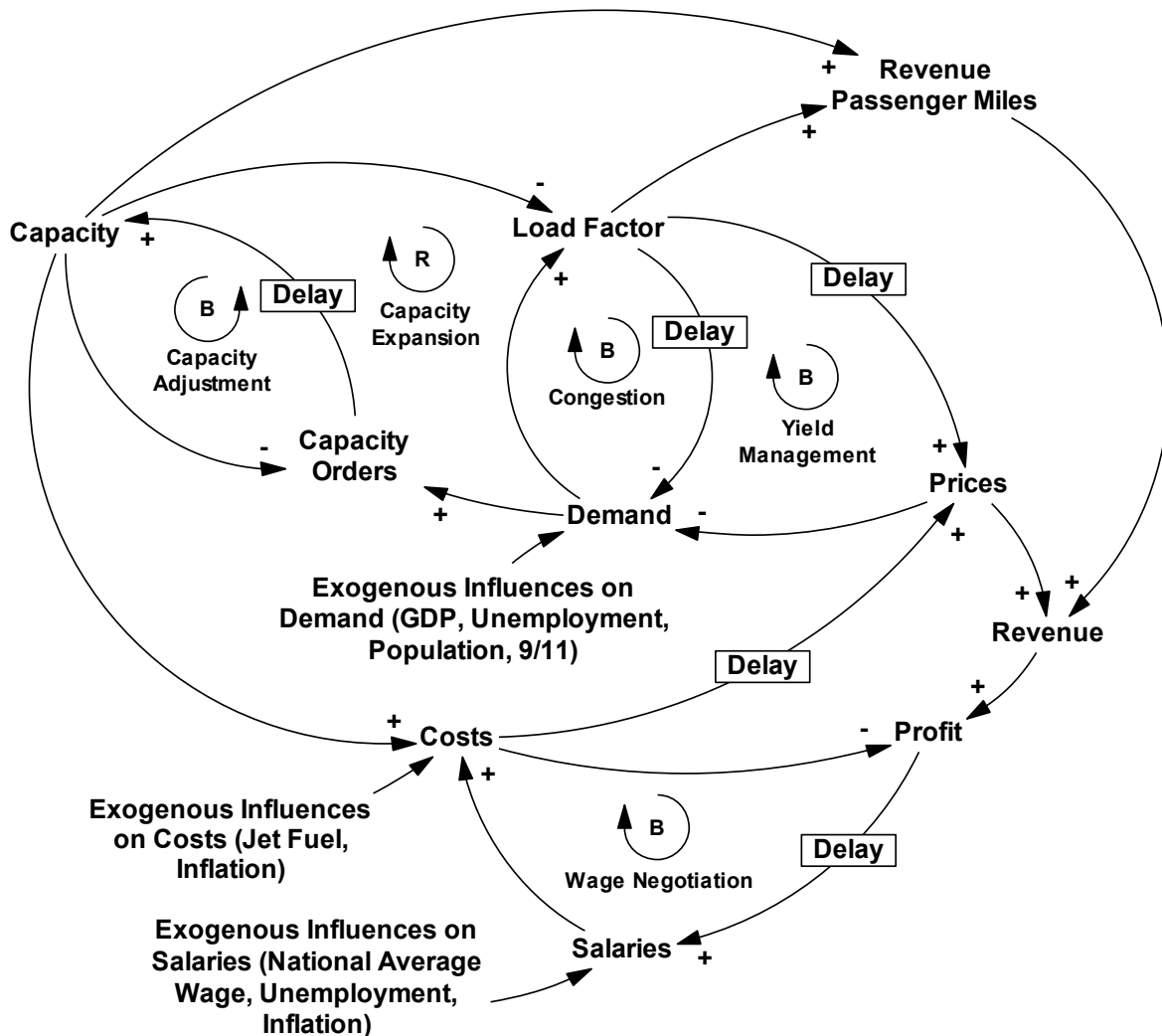


Figure 2: Overview of the model feedback structure and boundary.

Table 1 provides a summary of the model boundary, listing the main endogenous, exogenous and excluded variables.

Endogenous Variables	Exogenous Variables	Excluded Variables
Airline Capacity	Ancillary Fee Revenue	Advertising
Average Load Factor	Consumer Price Index	Aircraft Construction Capacity
Average Ticket Price	Employee Productivity	Aircraft Rental Costs
Average Wages for Airlines	Fuel Efficiency	Communication Costs
Cancellation of Capacity	Gross Domestic Product	Corporate Taxes
Capacity Ordering	Hours per Day Flown per Plane	Depreciation Expense
Cost per Available Seat Mile	Jet Fuel Price per Gallon	Food and Beverage Costs
Demand for Air Travel	Miles per Hour a Plane Travels	Insurance
Demand Forecasting	National Average Wage	Interest and Debt
Mothballing of Capacity	National Unemployment Rate	Landing Fees
Operating Profit	Normal Load Factor	Non-Aircraft Ownership Costs
Orders for Capacity	September 11 th Shock	Passenger Commissions
Reporting of Flow Variables	United States Population	Professional Services
Supply Line of Capacity	Yield Management Introduction	Utilities
Total Employment by Airlines		

Table 1: Model boundary diagram highlighting the most important endogenous, exogenous and excluded variables in the model. To the extent that excluded expenses vary with inflation they are indirectly represented in the model.

The model is organized into four principal sectors: Capacity, Demand, Prices and Costs. Here we describe the formulations for several critical variables. The online supplement (OS4) contains full model documentation using SDM-Doc (Martinez-Moyano, 2012) and all model, simulation, and experimentation documentation requirements (Rahmandad and Sterman, 2012).

Aggregate airline capacity is reported in available seat-miles per year. Each seat is assumed to fly a constant average number of miles per year determined from historical data for aircraft utilization. Airline capacity, the number of seats in the fleet, is modeled with a modified version of the standard stock control structure in the system dynamics literature (Sterman, 2000, Ch. 17). The stock of aircraft in service (Figure 3) is disaggregated into three vintages, with a mean aircraft lifetime of thirty years. The aircraft acquisition delay is assumed to be third order, with a mean acquisition time of two years (Airbus, 1998). Airlines are assumed to place orders to replace retirements of old aircraft and adjust capacity to demand given the normal load factor, while accounting for the supply line of aircraft on order, any returning to service from mothballing, and the expected rate of growth in demand (eq. 1 through 5):

$$\text{Orders} = \text{Max}(0, \text{DCA} + \text{SLA} + \text{SLA}_g - \text{RS}) \quad (1)$$

$$DCA = R + CA_g + (DC - C)/\tau_c \quad (2)$$

$$SLA = (DCA * \tau_m - SL)/\tau_s \quad (3)$$

$$SLA_g = S \cdot w \cdot g_e \quad (4)$$

$$CA_g = C \cdot w \cdot g_e \quad (5)$$

Aircraft orders are the sum of desired capacity acquisition (DCA), the supply line adjustment (SLA), and the two growth adjustments (CA_g and SLA_g), less capacity returning to service from the stock of mothballed aircraft (RS). DCA is the sum of retirements (R), CA_g , and a capacity adjustment based on the difference between desired capacity (DC) and current capacity (C). The strength of that capacity adjustment is controlled by τ_c , the estimated time to adjust capacity. Similarly, the supply line adjustment, SLA, is the gap between the desired and actual supply line, divided by the supply line adjustment time, τ_s . The desired supply line is determined, following Little's Law, by the product of the desired capacity adjustment, DCA, and the delay in manufacturing a plane (τ_m).

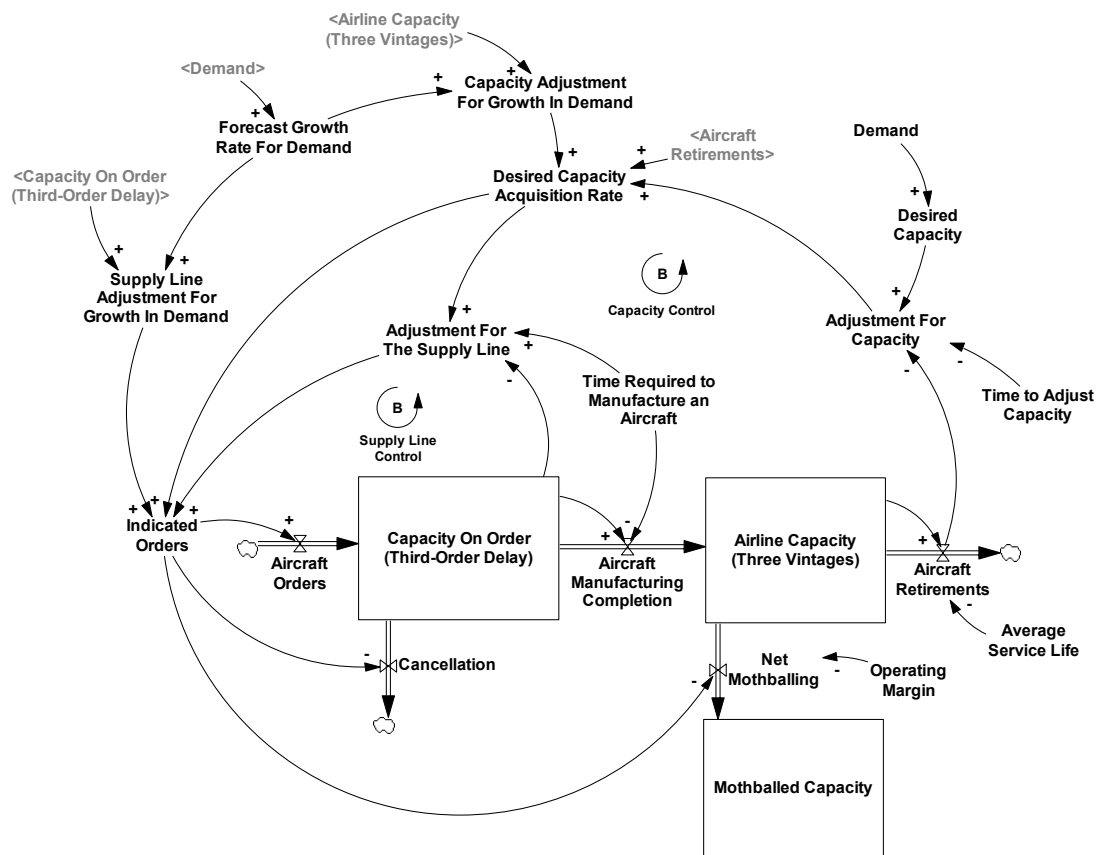


Figure 3: Overview of capacity and capacity acquisition.

We assume that airlines may plan for the growth of air travel demand. The growth adjustments, CA_g and SLA_g , increase orders based on g_e , the expected fractional growth rate in demand, with a weight, w , representing the extent to which the airlines actually account for the growth in demand when ordering capacity. The growth adjustments assure that under constant exponential growth there is no steady state error (if $w = 1$). A proof is available in the online supplement (OS2). The expected rate of growth is based on past growth rates using a standard trend function (Serman, 2000, ch. 16).

Demand for air travel is modeled as depending on population and air travel demand per capita. Population is exogenous. Per capita air travel demand depends on GDP per capita, the national unemployment rate, ticket prices, congestion, and an exogenous shock that captures the impact of the 9/11 terrorist attacks.

$$\text{Demand} = D_R \cdot \text{Pop} \cdot E_{GDP} \cdot E_{Unemp} \cdot E_{Price} \cdot E_{Cong} \cdot E_{9/11} \quad (6)$$

Air travel demand rises with growing incomes (Schafer, 1998), with an income elasticity S_{GDP} to be estimated:

$$E_{GDP} = \left(\frac{\text{GDP per capita}}{\text{Reference GDP per Capita}} \right)^{S_{GDP}} \quad (7)$$

Unemployment is a common independent variable in regressions used to forecast air travel demand, even when income effects are also included (Carson et al., 2011). We normalize unemployment by its historical average as shown below:

$$E_{Unemp} = \left(\frac{1 - \text{Unemployment Rate}}{1 - \text{Reference Unemployment}} \right)^{S_{UD}} \quad (8)$$

The unemployment rate is exogenous.

The effects of air travel prices and system congestion, measured by load factor, are:

$$E_{Price} = \left(\frac{\text{Price}}{\text{Price}_{Ref}} \right)^{S_{PD}} \quad (9)$$

$$E_{Cong} = \left(\text{smooth} \left(\frac{\text{Perceived Load Factor}}{\text{Normal Load Factor}} \right), \tau_{con} \right)^{S_{CD}} \quad (10)$$

S_{PD} is the price elasticity of demand, and reference price is the initial ticket price, scaled by inflation. The normal load factor has changed over the last 40 years with improvements in system operations and information technology. We model the reference load factor as

the best-fit quadratic for historical load factor³, and S_{CD} is the sensitivity of demand to congestion. There is a delay in the public's perception of congestion, so perceived load factor is modeled by first order smoothing. Since there is also a delay before congestion changes flying habits the ratio of perceived to normal load factor is smoothed again, with an adjustment time τ_{con} .

The terrorist attacks of September 11th 2001 immediately reduced air travel demand, with an effect that lingered for several years. The details of this formulation can be found in the online supplement (OS4).

Ticket prices are modeled with a standard price-discovery, hill-climbing formulation (Sterman, 2000, Ch. 13). Current ticket prices adjust with a delay to the indicated ticket price, which anchors on the current price and adjusts to pressures from profit margins, costs, and load factor:

$$Price = \int \frac{Price_{Ind} - Price}{\tau_p} + P_0 \quad (11)$$

$$Price_{Ind} = Price \cdot E_{Cost} \cdot E_{LF} \quad (12)$$

$$E_{Cost} = \frac{\text{Expected Passenger Cost} \cdot (1 + \text{Target Profit Margin})}{Price} \quad (13)$$

$$\begin{aligned} & \text{Expected Passenger Cost} \\ &= \frac{\text{Total Costs-Ancillary Fees}}{\text{Available Seat Miles} \cdot \text{Normal Load Factor}} \end{aligned} \quad (14)$$

$$E_{LF} = \left(\frac{\text{Load Factor}}{\text{Normal Load Factor}} \right)^{S_{SDP}} \quad (15)$$

Airlines in the model calculate their expected costs per passenger, on a seat mile basis, using current costs less any fees collected. Net cost is divided by the expected passenger volume, given by capacity and the normal load factor, to yield the expected cost per seat-mile, which is then marked up by the target profit margin. Total operating costs are the sum of costs from wages, costs from fuel, and other costs. Both fuel prices and fuel efficiency are exogenous. Other costs are modeled as an initial dollar amount per seat mile that grows with the Consumer Price Index.

Airline ticket prices also respond to imbalances between demand and supply, as indicated by load factor (Kimes, 1989). At the level of an individual carrier low load factors indicate that prices for the flight in question should fall. In the short term this will increase demand for that flight, and for the individual firm. Naturally, however, firm-level

³The quadratic approximation for normal load factor fits well over the period from 1970 to 2010, with an R^2 of 95.6%. The regression estimates are statistically significant at the 1% level. Omitting the quadratic term significantly degrades the endogenous model's fit for demand.

demand elasticity is much higher than industry-level demand elasticity (Oum et al., 1990), so most of the increase in the individual carrier's load factor comes at the expense of their rivals, who will respond with similar fare reductions. In the aggregate this causes prices to fall when load factors are low and rise when planes are relatively full. This relationship is captured in Equation 15.

While most yield management research is focused on pricing at the level of individual firms, in industry-level models such as the one developed here it is necessary to model the evolution of industry average prices, a common practice in system dynamics, including Meadows' (1970) commodity cycle model, Mass' (1975) business cycle model, Forrester's National Model (Forrester, 1979), many models of the oil industry (e.g., Davidsen et al., 1990), shipping industry (e.g., Randers and Göluke, 2007), electric utility industry (e.g. Ford, 1997), and others, including prior airline industry models (Liehr et al., 2001; Lyneis, 2000).

When yield management technology was introduced to the airline industry in 1985 ticket prices became much more responsive to load factor (Smith et al., 1992). To capture this effect the sensitivity of prices to the supply demand balance, S_{SDP} , is modeled as a step increase in 1985, the size of which is estimated during model calibration.

To model average airline employee wages we again employ a standard hill-climbing formulation in which wages respond to three pressures: profit margin, unemployment, and outside opportunities (wages in other industries). If there were no net effect from these pressures the average wage would increase with inflation.

$$Wage = \int \frac{Wage_{Ind} - Wage}{\tau_w} + W_0 \quad (16)$$

where τ_w is the delay in adjusting wages, and W_0 is the initial average wage.

$$Wage_{Ind} = Wage \cdot E_{Profit} \cdot E_{UnemW} \cdot E_{Opp} \cdot (1 + \Delta CPI) \quad (17)$$

Industry profitability is reported with a delay because it takes time for the parties in collective bargaining negotiations (airlines and unions) to form expectations about profits from past data. Wages tend to rise when airlines are relatively profitable and fall when they are less profitable:

$$E_{Profit} = \left(\frac{1 + Margin_{per}}{1 + Margin_{Ref}} \right)^{S_{MW}} \quad (18)$$

The reference margin in Equation 18 is the historical average margin for the industry, calculated from the data. The perceived margin, $Margin_{per}$, is modeled using first order exponential smoothing of operating profit margin, with a delay time to be estimated, along with the strength of the effect of profitability on wage negotiations, S_{MW} .

Wages ought to rise faster (slower) when unemployment is below (above) normal.

We model normal unemployment as the average historical value over the horizon of the model:

$$E_{UnemW} = \left(\frac{\text{Unemployment}}{\text{Normal Unemployment}} \right)^{S_{UW}} \quad (19)$$

Wages should also respond to wages in other industries. Since there is a skill premium offered for jobs in the airline industry, average airline wages are higher than the national mean. We assume airline wages respond to the national average wage (NAvgWage) adjusted by the average wage premium, with a sensitivity to be estimated.

$$E_{Opp} = \left(\frac{\text{Wage}}{\text{NAvg Wage} * \text{Wage Premium}} \right)^{S_{OW}} \quad (20)$$

Consistent with the literature in system dynamics (e.g. Sterman, 2000), and the broader literature in behavioral decision making and cognitive psychology (e.g. Stanovich, 2011), the decision rules for pricing, wages, aircraft orders, mothballing, etc., are boundedly-rational, behavioral heuristics, grounded in well-established evidence regarding the way managers make decisions in complex dynamic systems.

The online supplement (OS4) provides full documentation of the model.

Airline Industry Data

The data for parameter estimation come from the Air Transport Association (ATA), the nation's oldest and largest airline trade association, the Bureau of Transportation Statistics (BTS), and MIT's Airline Data Project (ADP).⁴ These data include available seat miles (capacity), revenue passenger miles (demand), average ticket price per revenue passenger mile (price), average wage per worker, including salary, benefits and other compensation (wage), and aggregate operating profit (profit). Data for U.S. population comes from the Census Bureau, while GDP data for the U.S. are from the Bureau of Economic Analysis and measured in real, year 2000 dollars per capita. The CPI, national average wage, and unemployment data come from the Bureau of Labor Statistics. Jet fuel prices per gallon and employee productivity are obtained from the ATA. Ancillary fees come from the ADP.

Parameter Estimation

We estimated model parameters by minimizing the weighted sum of the squared error between the model and the data simultaneously for each of the relevant data series:

⁴ATA: www.airlines.org; ADP: <http://web.mit.edu/airlinedata/www/default.html>. The ATA is now known as Airlines for America.

$$\min_{x \in \mathbb{R}} \sum_{i=0}^n \frac{\sum_{t=0}^{t_f} [(\hat{\theta}_{it} - \theta_{it})]^2}{\sqrt{\text{MSE}(\hat{\theta}_i)}} \quad (21)$$

where the data series θ_i include historical demand, prices, wages, operating profit, etc., depending on the particular partial model test or full model estimation performed. The error from each series is weighted by the root mean square error of the model estimate from the previous calibration run. The process is iterated until the weights and estimates converge.

The sum of squared errors for each variable included in the estimation process is weighted by the reciprocal of the root mean squared error between the simulated and actual data series. Doing so assures, assuming normally distributed errors, that the total estimation error will be distributed chi-square, allowing us to estimate confidence intervals for each parameter using a Markov chain Monte Carlo (MCMC) method (Gelman et al., 2003). We use MCMC to simulate the distribution of the log likelihood payoff surface given joint changes in the parameters. The MCMC algorithm was implemented using commercially available software and we provide a detailed description in the appendix (A2). Convergence took approximately 1.2 million model runs, or close to 16 hours of desktop computer time.

Partial model testing (Homer, 2012) was the first stage of our parameter estimation process. Each sector of the model was isolated and driven by historical data for the inputs to that sector. In the partial model test of the demand formulation (eq. 6), we use historical ticket prices and load factor rather than their endogenous values, along with historical GDP, unemployment, and population, to estimate demand. The partial model test for growth expectations uses historic demand to fit the trend function for expected growth in demand (an input to the capacity decision) against ten years of FAA demand forecasts. The partial model test of industry capacity replaced endogenous demand and profit with historical demand and operating profit. The partial model test for costs used historical wages and capacity together with exogenous fuel costs, efficiency, and inflation. The partial model test for industry wages used historical operating profit along with national unemployment, average wage, and inflation. The partial model test for price setting used historical operating costs, demand and capacity instead of their endogenous formulations.

The estimated parameters from partial model testing are reported in Table 2, along with the 95% confidence intervals estimated by the MCMC method. Figure 4 compares the simulated and actual data for the partial model tests, and Table 3 reports goodness of fit measures. Overall the partial model tests have low error as a percentage of the mean and low bias, as shown by the Theil inequality statistics, indicating the errors are generally unsystematic.

The estimated parameters in the partial model tests are reasonable. The structure for the impact of the 9/11 terrorist attacks captures an immediate decline in air travel,

and the subsequent reduction in demand due to fear and the resulting security measures, which is assumed to gradually decrease over time. The estimated parameters suggest an immediate drop of nearly 15% in demand and a decay time of approximately 9 years. Sensitivity tests involving first order delays, higher order delays, and other specifications for the effect of 9/11 on demand all showed time constants on the order of the one reported here. The long decay time suggests the impacts of 9/11 have been persistent, perhaps a result of later, failed attacks such as the shoe and underwear bombers, or the inconvenience and costs of the security measures implemented since 2001. Alternatively, it is possible that some other factors caused a shift in the demand for air travel after 2001.

Parameter Name	Eq. #	Lower Bound of 95% CI	Partial Model Estimate	Upper Bound of 95% CI
Capacity				
Time to Adjust Capacity (years)	2	0.124	0.132	0.1452
Supply Line Adjustment Time (years)	3	0.083	0.100	0.1106
Weight on Demand Forecast Orders (fraction)	4,5	0.554	0.683	0.8625
Demand				
Reference per Capita Demand (seat*miles/year)	6	1039	1044	1047
Income Elasticity of Demand (dmnl)	7	1.01	1.12	1.19
Price Elasticity of Demand (dmnl)	9	-0.481	-0.406	-0.351
Sensitivity of Demand to Congestion (dmnl)	10	-0.524	-0.472	-0.404
Congestion Adjustment Time (years)	10	1.49	1.76	1.86
Strength of Unemployment Effect on Demand (dmnl)	8	1.90	1.93	2.00
Size of 9/11 Effect (fraction)	OS	0.129	0.146	0.164
Public Perception of Terrorism Decay Time (years)	OS	8.39	8.91	9.46
Price and Unit Costs				
Initial Other Variable Costs (dollars/(seat*mile))	OS	0.0190	0.0193	0.0195
Time to Adjust Ticket Prices (years) ⁵	11	0.083	0.083	0.130
Target Profit per Passenger (dollars/(seat*mile))	13	0.0274	0.0332	0.0393
Effect of Yield Management on the Sensitivity of Price to Demand Supply Balance (dmnl)	15	2.57	3.02	3.48
Base Sensitivity of Price to Demand Supply Balance (dmnl)	15	0	0	0.172
Salary				
Time to Change Worker Compensation (years)	16	1.06	1.07	1.08
Strength of Unemployment Effect on Wages (dmnl)	19	-0.0034	-0.0003	0
Strength of Margin on Worker Compensation (dmnl)	18	0.291	0.372	0.409
Strength of Outside Opportunities on Worker Compensation (dmnl)	20	0	0.0007	0.0092
Margin Perception Delay (years)	OS	3.11	3.24	3.48

Table 2: Estimated parameters from partial model testing, with Markov chain Monte Carlo 95% confidence intervals. The equation number "OS" indicates that the equation is reported in the online supplement (OS4), not in the paper.

⁵ The lower bound for all time constants was 0.083 years, approximately one month.

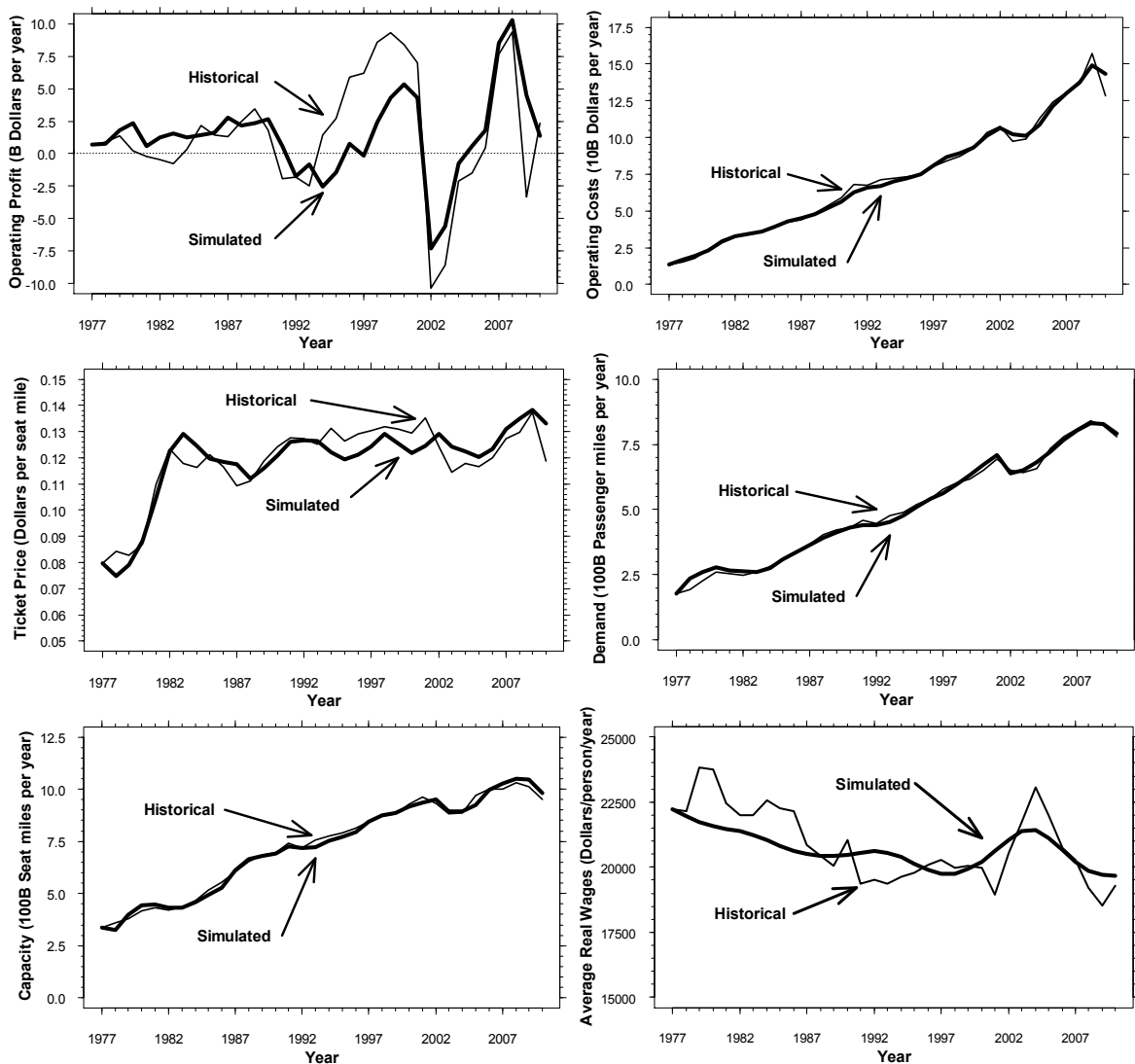


Figure 4: Partial model test results plotted against historical data.

Variable	R^2	MAE/M	RMSE/M	U^M	U^S	U^C
Capacity	99.3%	.0211	.0265	.0159	.1799	.8190
Demand	99.4%	.0248	.0315	.0056	.0594	.9350
Wages	43.5%	.0401	.0502	.0054	.5137	.4809
Cost	99.6%	.0260	.0368	.0708	.0574	.8718
Prices	86.4%	.0398	.0481	.0079	.0075	.9846
Profit	56.4%	N/A	N/A	.0011	.1799	.8190

Table 3: Partial model fits to historical data for 1977-2010. R^2 is defined as one minus the ratio of the sum of the squared error to the total sum of squares. MAE/M is mean absolute error divided by the mean of the data. RMSE/M is the root mean square error divided by the mean of the data. U^M , U^S , and U^C are the Theil inequality statistics (Sterman, 2000, ch. 21), which partition the MSE into the fraction arising from bias (unequal means of simulated and actual data), unequal variances, and unequal covariation, respectively. MAE/M and RMSE/M are not reported for profit because average historical profit is very close to 0.

The partial model tests indicate that the model reproduces sector-level behavior quite well, with the exception of the average airline industry wage. The fit of the model to the wage data is somewhat lower than the fit to the other variables. However, the mean absolute error is only 4% of the average of the historical wage data and the bias is very small. The fit of the model to the data, including the fit for wages, compares favorably against other models in the system dynamics literature and in related modeling traditions such as the forecasting literature. For example, Makridakis et al. (1982) examined the performance of a wide range of forecasting and modeling methods, using data from a large variety of systems. Typical calibration errors (assessed by the mean absolute percentage error, MAPE), for a subsample of 111 data series, were about 20% for non-seasonal methods applied to the raw data, about 11% for methods that accounted for seasonal adjustments, and about 9% for the non-seasonal methods applied to the seasonally adjusted data.

Nevertheless, additional research into the determinants of airline wages would help to address the source of the unexplained variation in airline wages and whether these sources are plausibly endogenous or reflect factors unrelated to the cycle in aggregate profitability. For example, industry wages may be heavily influenced by bankruptcies of individual carriers and labor actions such as strikes, both of which are difficult to predict and not modeled here.

The partial model tests examine the ability of individual formulations to replicate industry dynamics given the actual, realized values of the inputs to each formulation or decision. However, the partial model tests cut important feedbacks in the system, so it is also necessary to examine the ability of the full, endogenous model to fit the data.

Full model estimation results (Tables 4, 5; Figure 5) improve the fit for demand, price, and operating profit compared to the partial model results. The fit for the other variables remains similar. All series show low bias and, with the exception of wages, low unequal variation. The estimated parameters are plausible and the MCMC confidence bounds generally tight. The estimated values of a number of parameters are very similar to the values in the partial model tests, for example, the size and decay time of the 9/11 effect. Several others, however, differ from the partial model estimates.

In particular, in the full system estimation the capacity sector of the model became significantly less reactive, with longer time constants for capacity and supply line adjustment, and a smaller response to demand forecasts. In the partial model test for capacity acquisition the time constant controlling the adjustment for the supply line was 0.1 years, suggesting that airlines are keenly aware of and swiftly adjust the supply line of aircraft on order as the desired number of aircraft they seek to acquire changes. Evidence from experimental studies (e.g. Sterman, 1989; Aramburo et al., 2012; Croson et al., forthcoming), and from other industries (e.g., commercial real estate and shipbuilding, see Sterman, 2000; Randers and Göluke, 2007) suggests weak supply line adjustment and a role for inadequate supply line control in the genesis of industry cycles. However, the high price of aircraft, concentrated nature of the industry, and contractual terms for aircraft orders may favor fully accounting for the supply line. The supply line adjustment time in

the full model estimation is longer and more plausible, though at about 4 months, still short enough to suggest that airlines are quite sensitive to the supply line of capacity on order. Exploring this issue further would require data on order cancellations, aircraft completion, and the supply line of planes, perhaps at the level of individual manufacturers, data that are not publicly available.

Parameter	Partial Model Estimate	Eq.	Lower Bound of 95% CI	Full Model Estimate	Upper Bound of 95% CI
Capacity					
Time to Adjust Capacity (years)	0.132*	2	0.459	0.476	0.490
Supply Line Adjustment Time (years)	0.100*	3	0.308	0.372	0.388
Weight on Demand Forecast Orders (fraction)	0.683*	4,5	0.173	0.211	0.242
Demand					
Reference per Capita Demand (seat*miles/year)	1044*	6	1145	1146	1162
Income Elasticity of Demand (dmnl)	1.12	7	1	1.01	1.03
Price Elasticity of Demand (dmnl)	-0.406*	9	-0.333	-0.325	-0.289
Sensitivity of Demand to Congestion (dmnl)	-0.472*	10	-3.87	-3.01	-2.99
Congestion Adjustment Time (years)	1.76	10	1.24	1.36	1.59
Strength of Unemployment Effect on Demand (dmnl)	1.93*	8	2.96	3.06	3.04
Size of 9/11 Effect (fraction)	0.146	OS	0.158	0.163	0.171
9/11 Impact Decay Time (years)	8.91	OS	8.88	8.99	9.43
Price and Unit Costs					
Initial Other Variable Costs (dollars/(seat*mile))	0.0193*	OS	0.0163	0.0187	0.0189
Time to Adjust Ticket Prices (years)	0.083*	11	0.132	0.222	0.271
Target Profit per Passenger (dollars/(seat*mile))	0.0332*	13	0.0052	0.0112	0.0166
Effect of Yield Management on the Sensitivity of Price to Demand Supply Balance (dmnl)	3.02	15	3.44	3.78	3.802
Base Sensitivity of Price to Demand Supply Balance (dmnl)	0	15	0	0	0.033
Salary					
Time to Change Worker Compensation (years)	1.07	16	1.08	1.10	1.11
Strength of Unemployment Effect on Wages (dmnl)	-0.0003	19	-0.0079	-0.0007	0
Strength of Margin Effect on Worker Compensation (dmnl)	0.372*	18	0.073	0.116	0.131
Strength of Outside Opportunities Effect on Worker Compensation (dmnl)	0.0007	20	0	0	0.0047
Margin Perception Delay (years)	3.24*	OS	3.60	3.68	6.45

Table 4: Estimated parameters from full model results, with Markov chain Monte Carlo 95% confidence intervals, and partial model parameters for comparison. Partial model estimates marked with an asterisk (*) are statistically significantly different from the full model estimates at the 5% level.

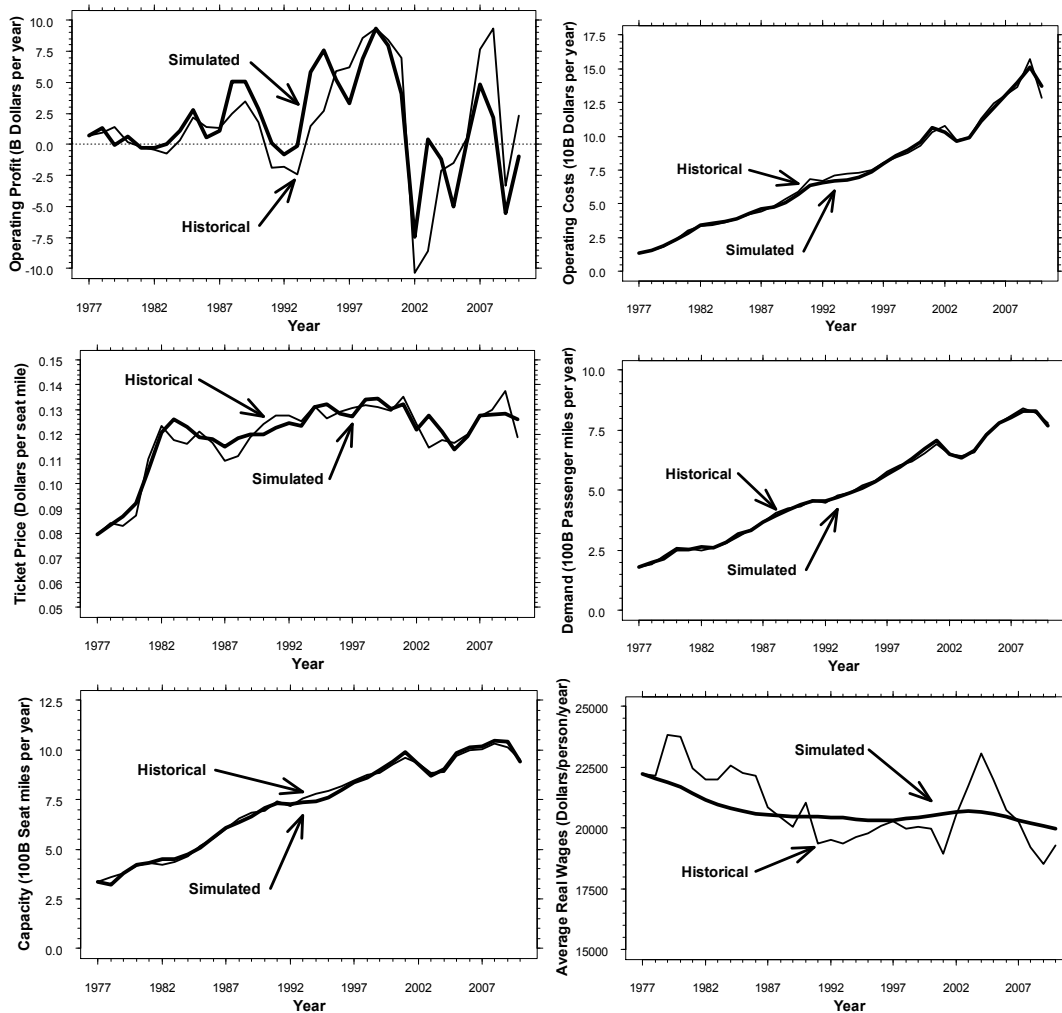


Figure 5: Full model results plotted against the historical data.

Variable	R^2	MAE/M	RMSE/M	U^M	U^S	U^S
Capacity	99.4%	.0207	.0249	.0011	.0557	.9432
Demand	99.8%	.0148	.0179	.0010	.0008	.9981
Wages	50.83%	.0407	.0497	.0098	.6257	.3645
Cost	99.6%	.0278	.0360	.0852	.0331	.8818
Prices	90.76%	.0300	.0384	.0176	.0304	.9520
Profit	62.8%	N/A	N/A	.0085	.0894	.9021

Table 5: Goodness of fit for full model, 1977-2010.

What accounts for the differences in parameter estimates between the partial and full models? First, the payoffs are different: in the partial model tests, the payoff is the fit to the focal variable in each sector: demand for the demand sector, capacity for the capacity sector, total cost for the cost sector and so on. In the full model estimation, the likelihood function is the sum of squared errors for all the key variables, specifically, demand,

capacity, prices, profits, and average wages, weighted by $1/\text{RMSE}$ for each. Second, the likelihood function for the full model appears to have a flat optimum. Over one million MCMC runs were needed to arrive at stable estimates for the confidence bounds. Further, to prevent convergence to local optima we used multiple starting points in the parameter space. Many of these restarts discovered unique local maxima, indicating that the global likelihood surface is relatively flat over the range of plausible values. Recent work on parameter testing and model validation (Hadjis, 2011; Groesser and Schwaninger, 2012) use relatively simple models to advocate for particular approaches to parameter identification, estimation and model testing. The airline industry context however, like many policy relevant settings, involves common and troublesome issues arising from endogeneity, collinearity, under-identification, and flat optima, rendering these approaches potentially problematic and indicating a need for more research.

Model Analysis

Oscillations in dynamic systems arise from negative feedbacks with significant phase lag elements (time delays). System dynamics models of earnings cyclicality have found that delays in the negative feedbacks controlling inventory, capacity acquisition or other resources are the underlying causes of cyclical movements in the economy and for many industries and commodities (e.g., Meadows, 1969; Chen et al., 2000; Sterman, 2000, chs. 17, 19 and 20; and Randers and Goluke, 2007). Unsurprisingly, our results are consistent with this mechanism: delays in the negative feedbacks regulating airline industry capacity as demand and profitability change contribute to the oscillation observed in industry profitability. However, many prior studies find that the amplitude and persistence of industry cycles are increased by the failure of industry participants to account sufficiently for the supply line of capacity on order. The failure to account for the supply line is well supported by experimental, econometric, and field evidence (e.g., Sterman, 1989; Sterman, 2000, Ch. 17; Randers and Goluke, 2007), and previous models of the airline industry (Liehr et al., 2001) also highlight the role of the supply line in profit instability.

However, supply line adjustment is only one of many delayed negative feedbacks in the airline industry. Our estimation results provide little evidence for failure to account for the supply line of aircraft on order as a source of the cycle in airline industry profitability. If industry participants, particularly the aircraft manufacturers, were unresponsive to the supply line of unfilled orders, then the estimated time constant for supply line adjustment would be very long, and longer than the capacity adjustment time. Instead, the supply line adjustment time we estimate is about the same as the capacity adjustment time in both the partial and full model tests. The result is plausible compared to, say, the real estate industry, where evidence suggests very low salience and responsiveness to the supply line (Sterman, 2000, Ch. 17.4.3). The real estate market is characterized by many producers, low barriers to entry, and therefore low experience among developers, and heterogeneity in building location, quality and price. It also is difficult to measure the supply line in real estate since it includes potential projects and

projects in various stages of permitting and financing, not only those under construction, and these projects differ by location, size, and other attributes that lower their comparability.

In contrast, the airline market is characterized by a small numbers of producers, high barriers to entry, and a small number of product variants. The supply line of capacity on order is well known to both manufacturers and their customers. These conditions favor a more complete accounting for the supply line in ordering decisions.

While our model suggests that airlines and manufacturers are unlikely to underweight the supply line in ordering decisions substantially, other feedbacks and time delays cannot be accounted for so easily. The role of these compensating feedbacks in the genesis of the cycle can be illuminated using the response of the fully endogenous model to a 1% step increase in population from an initial equilibrium⁶ (Table 6, Figure 6).

Test	Base Case: Supply Line Adjustment (SLA) Time: 0.372 years	Supply Line Adjustment (SLA) Time: 0.083 years	SLA Time: 1 year	SLA Time: 10 ⁹ years
Percent Undershoot	23.5%	13.2%	26.6%	20.1%
10% Settling Time	2.67 years	2.16 years	3.27 years	6 years
Damping Ratio	0.419	0.542	0.386	0.455
Oscillation Period	3 years	2.6 years	3.6 years	8 years

Table 6: Step response tests of the model. Operating profit is the output in each case. Undershoot is measured relative to the steady state value of profit at the end of the model run. The damping ratio (DR) was calculated by treating the model as a second-order system so that $DR = \sqrt{(\ln \%U)^2 / [\pi^2 + (\ln \%U)^2]}$, where %U is the percent undershoot. (Brown, 2007)

In the base case using the full model parameter estimates a step change in population causes an oscillatory response of operating profit, with a roughly 3 year period, settling time of 2.67 years and damping ratio of about 0.4 years. When the supply line adjustment time is increased to 1 year, three times longer than the estimated capacity adjustment time, the cycle period extends to 3.6 years, the settling time lengthens by about 0.6 years, and the damping ratio falls slightly, as seen in Table 6. Fully disabling the supply line adjustment feedback by setting the adjustment time to an essentially infinite value (one billion years) lengthens the cycle period further, to about 8 years, and lengthens the settling time, while increasing damping compared to the base case. Similarly, dramatically shortening the supply line adjustment time, to 1 month, shortens the cycle period by 0.4 years, cuts the settling time by about 6 months, and increases the damping ratio, but the cycle is not eliminated. The results show that the extent to which airlines account for the

⁶ All exogenous time series were set to their initial values and initial conditions for the state variables were set to start the model in dynamic equilibrium.

supply line of capacity on order matters to stability, but also that the oscillation in airline profitability is not solely created by the failure of industry actors to account for the supply line. Other negative feedbacks in the model, such as yield management, congestion, and capacity adjustment all contribute to the oscillation regardless of the strength of the supply line adjustment loop.

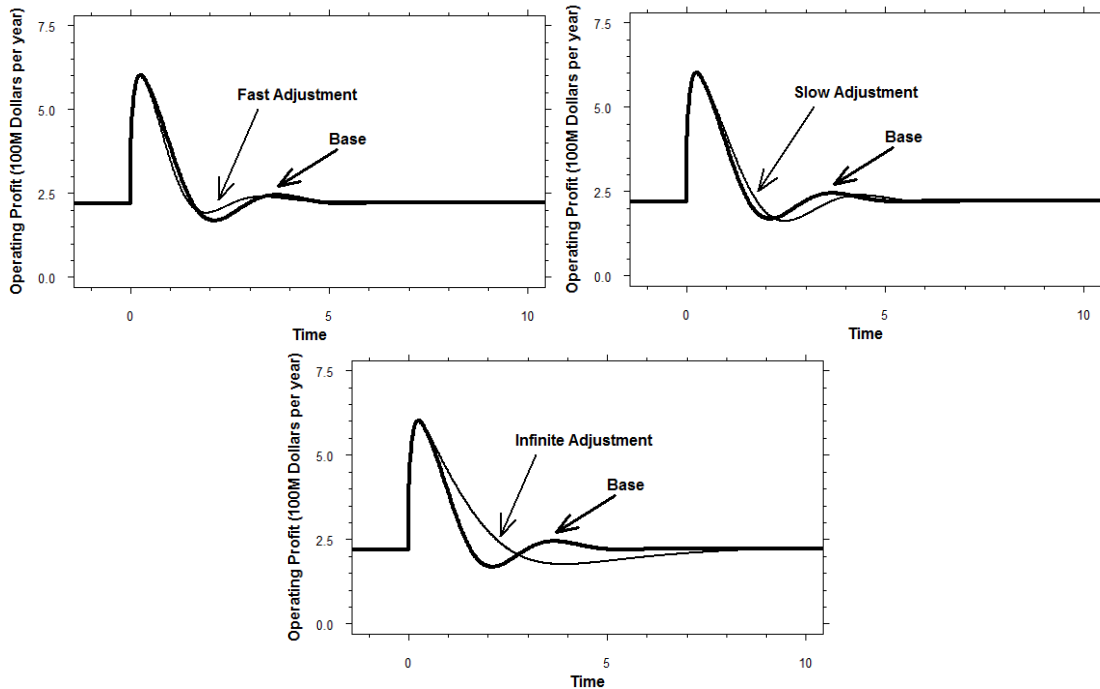


Figure 6: The response of operating profit to a 1% step increase in demand. The base case uses the parameters estimated during full model calibration, the infinite adjustment case sets the supply line adjustment time to 10^9 years, the slow adjustment case sets the supply line adjustment time to 1 year, and the fast adjustment case sets the supply line adjustment time to 0.083 years.

Interestingly, yield management is a particularly important determinant of the stability of profit in our model. The yield management feedback acts when increases in demand cause higher load factors, raising average industry ticket prices, which then decrease demand in a negative feedback.

The step responses reported in Table 7 and Figure 7 show how dramatically varying the sensitivity of price to the demand supply balance (Eq. 15) influences system stability and the variability of both profit and capacity. Eliminating yield management from the price setting heuristic worsens every measure of system stability; whereas doubling the sensitivity of price to the demand supply balance increases the stability of the system substantially.

Test	Base Case	No Yield Management ($S_{SDP} = 0$)	More Yield Management ($S_{SDP} = 2 \times \text{base case}$)
Percent Undershoot	23.5%	45.8%	1.2%
10% Settling Time	2.67 years	7.3 years	2.3 years
Damping Ratio	0.419	0.241	0.816
Oscillation Period	3 years	2.6 years	N/A
Operational Leverage ⁷	172%	87%	209%

Table 7: Step response tests of the model varying the sensitivity of price to the demand supply balance (Eq. 15). Operational leverage is calculated by determining percentage change in profit between equilibrium and the first peak in the step response. That quantity is then divided by the percentage change in demand (1%).

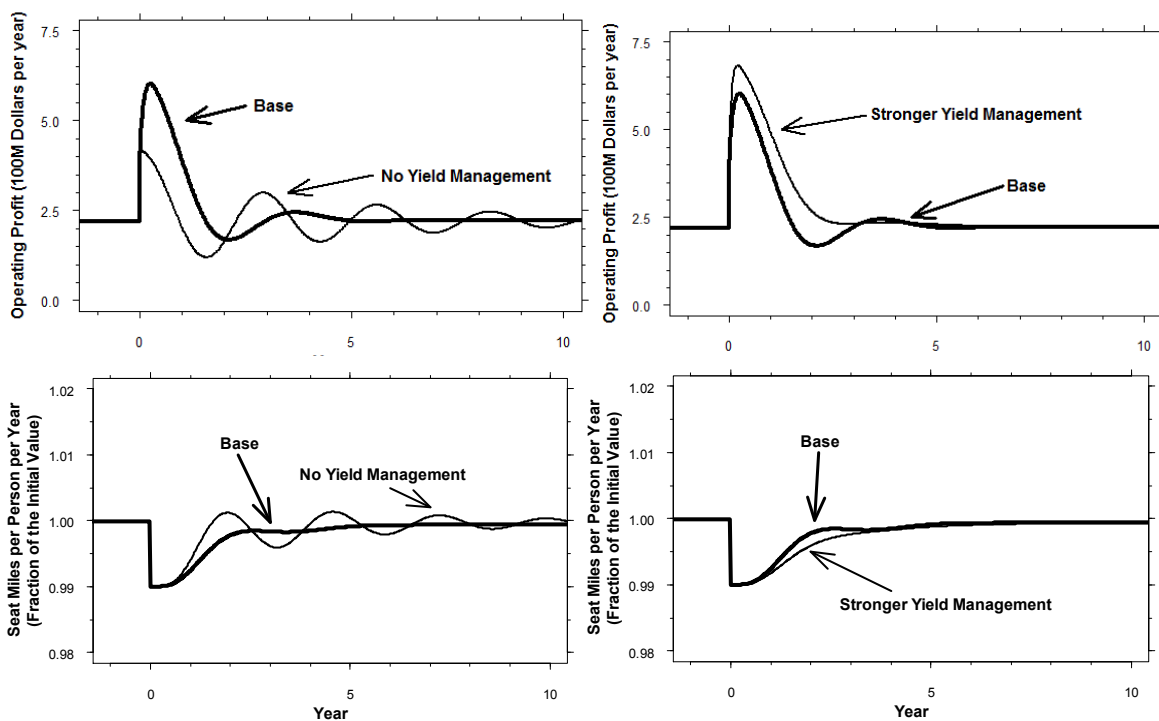


Figure 7: Step response of the model varying the sensitivity of price to the demand supply balance (yield management). Top: Impact on operating profit. Bottom: Impact on capacity (seat miles of capacity per capita).

The logic behind this result is straightforward. Because the estimated time to adjust ticket prices (eq. 13) is very short (0.083 years, our lower bound for time constants) compared to the lags in capacity acquisition, yield management acts as an effectively first-order negative loop that damps the oscillatory response of capacity and other variables to demand shocks. The stronger the effect of load factors on price, the greater the

⁷ Operational leverage is measured using the peak instantaneous value observed in our step response tests.

stabilizing effect of the price-demand feedback. Consider an analogy to the mass-spring-damper system. The damping force is proportional to the velocity of the mass, completing a first-order negative feedback around velocity (higher velocity, more opposing force, lower acceleration, lower velocity) that attenuates the amplitude of the oscillation. Stronger damping increases the stability of the system.

Yield management works in much the same way. With strong yield management, average airline ticket prices will rise much more quickly than capacity when there is a positive demand shock. Because demand for air travel is elastic this increase in price works to oppose the change in demand. The stronger this demand “friction” the more damped the system will be.

The stabilizing influence of yield management can also be seen in the step response of capacity. When demand increases at time zero the adequacy of capacity falls. In the base case capacity slowly recovers, with a slight overshoot and oscillation, to its equilibrium value. Increasing the strength of yield management slows this approach to equilibrium and eliminates the overshoot, while removing yield management dramatically reduces damping.

Thus, stronger yield management improves system stability by increasing damping. However, traditional measures of system stability are not the only metrics that matter to the airlines and other stakeholders. Stronger yield management increases damping but also increases the magnitude of the change in operating profit resulting from the demand shock (Figure 7). The response of profit to changes in demand is known as operational leverage in managerial accounting and is an important indicator of risk.

Accountants use the ratio of the fixed and variable costs of an enterprise to calculate operational leverage. A higher ratio of fixed to total costs implies higher operational leverage, since only variable costs change with the number of units sold and therefore the jump in demand will cause revenue to increase by much more than costs. Higher operating leverage indicates higher inherent risk, since a decrease in demand under high operational leverage reduces costs much less than it reduces revenue.

Recent research has suggested that high operational leverage can justify implementation of revenue management⁸ for firms (Huefner and Largay, 2008). The argument is that revenue management, by increasing unit sales, will have a greater impact on profit if operational leverage is high because incremental revenue contributes more to the bottom line.

However, such arguments typically don’t consider the impact of revenue (yield) management on the volatility of profits.

Consider an unanticipated, positive demand shock. Profits rise as load factor rises. If price also rises in response to the increase in demand (and if the aggregate demand

⁸ Revenue management is the more general term for yield management. The word “yield” is used primarily in the airline industry.

elasticity is less than one⁹) then profits will increase even further, because each seat is sold at a higher average price. Hence, the stronger the effect of yield management on prices, the greater the operational leverage of the industry. Figure 7 shows that stronger yield management stabilizes the fluctuations in capacity and increases damping, but the initial response of profit to the demand shock is much larger. So, while yield management increases the damping of the system, it simultaneously increases the short run volatility of profits, and hence the risk investors face, as the industry responds to demand shocks.

To explore the relationship between yield management, operational leverage, and risk more thoroughly, Figure 8 maps average profit as it depends on the strength of the two factors affecting price in the model: the sensitivity of price to costs (markup) and the sensitivity of price to load factors (yield management). The height of the surface is average profit over each model run.

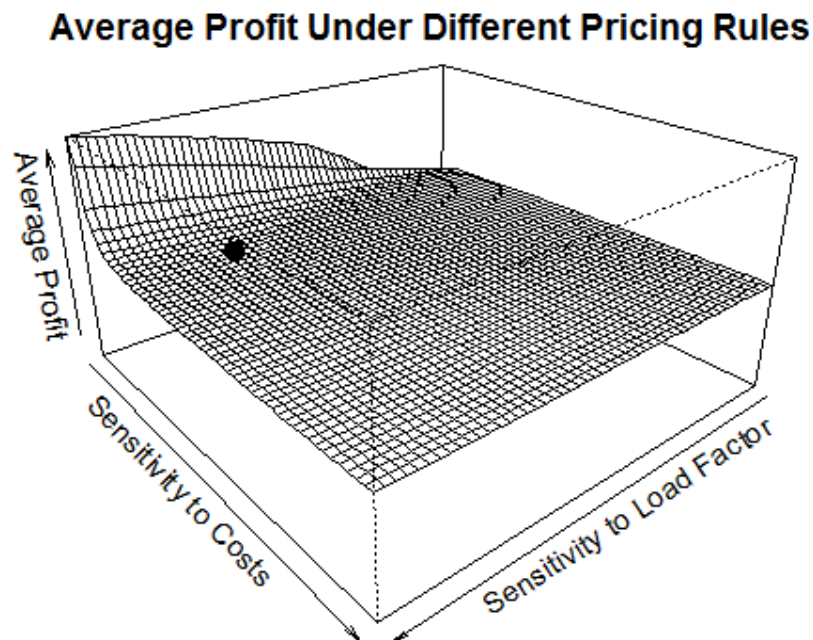


Figure 8: Average annual operating profit between 1977 and 2010 as a function of the sensitivity of prices to load factor and to costs. The black circle indicates the estimated parameters for the full model.

Over most of the surface, including the neighborhood of the estimated parameters, the gradient indicates that higher sensitivity to load factor and lower sensitivity to costs raises average profit. That is, more aggressive yield management boosts average profitability. This suggests one reason why the industry has steadily evolved towards greater reliance on the use of yield management, including more categories of fares, more frequent fare

⁹ The estimated elasticity of air travel demand with respect to price is much less than 1 (about 0.32 in the full model), and other studies find similarly low values (Oum, Waters, and Yong, 1990).

changes, and more sophisticated models to predict future demand during the reservations window (Belobaba, 1987). As shown in Figure 9, greater reliance on, and more effective, yield management technology has enabled the average load factor of the fleet to rise steadily, from about 0.6 in the 1980s to more than 0.8 in the last decade.

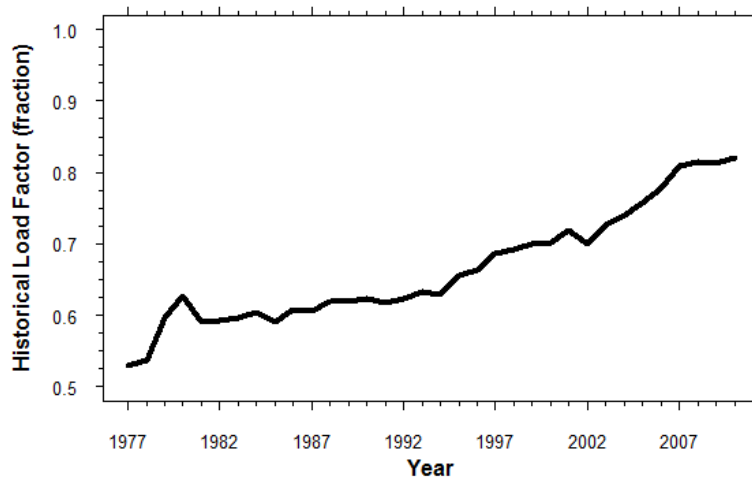


Figure 9: Average load factor since deregulation has climbed steadily, aided by better technology for reservations and increasing use of yield management to balance demand with available capacity.

However, while aggressively cutting prices to fill empty seats boosts average profit, it also increases the response of profit to demand shocks and other perturbations. Figure 10 shows the standard deviation of profit over the same parameter space shown in Figure 8.

Pricing policies that generate higher average profits in Figure 8 also induce higher variability in profits in Figure 10, suggesting that the benefits of yield management are less clear: while higher average profits are obviously desirable, greater variability in profits increases the risk premium investors demand, makes bankruptcy more likely during industry downturns, and exacerbates demands for salary increases during boom periods and the likelihood of layoffs and labor problems during downturns.

To explore the risk-return tradeoff inherent in the implementation of yield management, Figure 11 shows average profit divided by the standard deviation of profit, analogous to the Sharpe ratio (Sharpe, 1994). The risk-adjusted return surface suggests that the industry's current pricing policy, as indicated by the full model estimates, which capture the extensive use of yield management, may be close to the "worst of both worlds" by delivering higher profit volatility than a policy that prices based on unit cost, but lower average profit than a policy that exclusively uses yield management.

Of course what's best for the industry is not necessarily what's best for individual carriers. The game theoretic and competitive issues here are important, but beyond the scope of the present model and paper. Future work should consider disaggregating the

Standard Deviation of Profit Under Different Pricing Rules

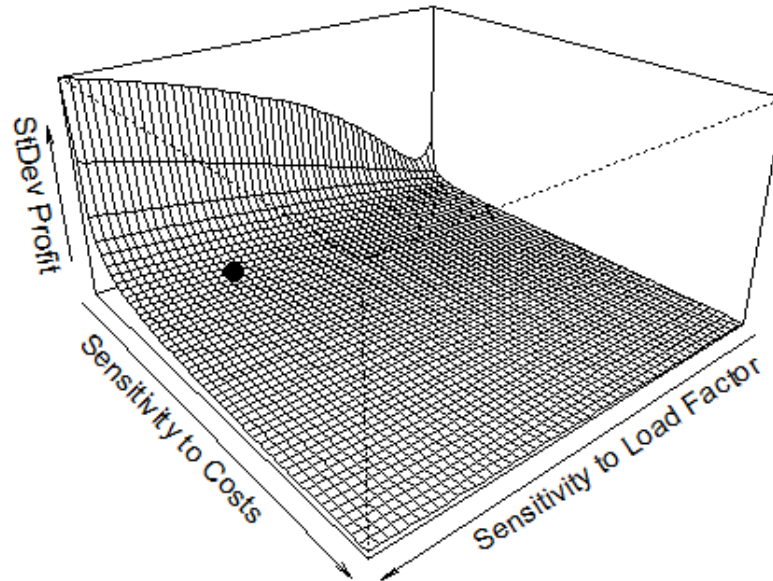


Figure 10: Standard deviation of annual operating profit between 1977 and 2010 as a function of the sensitivity of prices to load factor and to costs. The black circle indicates the estimated parameters for the full model.

Scaled Profit Under Different Pricing Rules

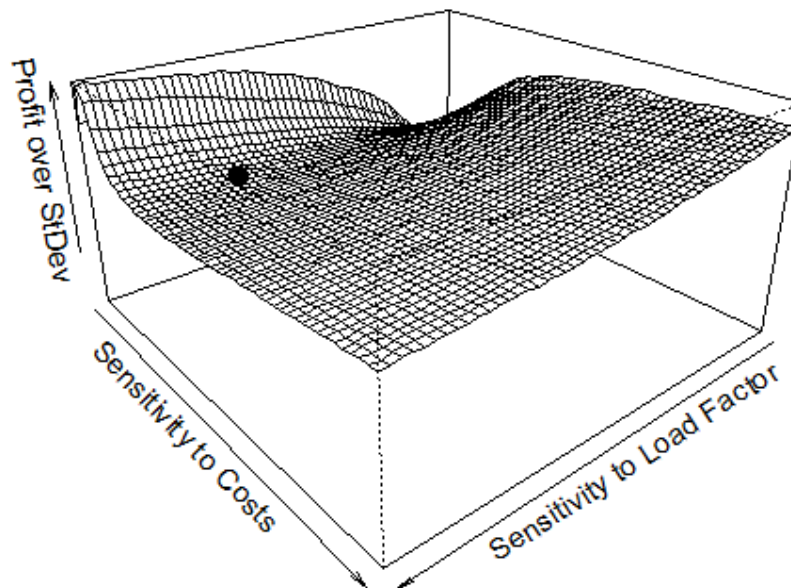


Figure 11: Risk-adjusted return, given by the Sharpe ratio (average profit divided by the standard deviation of profit), between 1977 and 2010 as a function of the sensitivity of prices to load factor and to costs. The black circle indicates the estimated parameters for the full model.

model to represent competition among carriers, including price setting, entry, exit, bankruptcy, and interactions between the carriers and financial markets that supply needed capital. Nevertheless, our results suggest that the airline industry may currently be decreasing risk-adjusted profit as an unintended consequence of the effort to boost profitability by filling otherwise empty seats.

Limitations and Extensions

The model, like all models, could be extended and improved. As discussed above, one possibility is to disaggregate to the level of individual airlines to examine the competitive dynamics that take place in the context of the overall industry cycle. Another concern is the period of the profit cycle we measure. The observed period of airline profit cycles is on the order of 10 years (Hansman and Jiang, 2005), yet, with the best-fit parameters the step response of the model exhibits a period of approximately 3 years.

Of course the period of the cycle observed in the data cannot be compared directly to the period of the step response. The observed period is the response of the industry to perturbations spanning a wide range of frequencies, from short-term noise to longer-term cyclical movements in demand induced by the business cycle to even slower changes in demographics, airport capacity, and other determinants of air travel demand. The behavior of a dynamic system responding to shocks is the convolution of the closed-loop frequency response of the system with the power spectrum of the noise inputs. Because there is significant power in the low frequency components of the perturbations in demand, the observed cycle period will be longer than the period observed in the step response. Given that the full model fits the data well with plausible parameters, there is no evidence to suggest that the period of the oscillatory response to the idealized step input is problematic. However, future research should explore this issue further.

A related issue concerns aircraft manufacturing capacity. We have modeled the delay in acquiring new aircraft as a constant, implicitly assuming that aircraft manufacturing is uncapacitated. In reality, aircraft manufacturing capacity can constrain the delivery of new aircraft, lengthening the aircraft acquisition delay and potentially increasing the natural period of the endogenous industry cycle. The online supplement (OS3.2) reports a structural sensitivity test that adds manufacturing capacity constraints and endogenous manufacturing capacity to the model. Under certain parameterizations (long delays in the response of manufacturing capacity to changes in aircraft orders), the inclusion of capacitated deliveries lengthens the period of the profit cycle observed in the step response. Importantly, however, the inclusion of endogenous manufacturing capacity does not change our findings concerning yield management: even with endogenous manufacturing capacity constraints, stronger yield management helps stabilize the capacity cycle but at the expense of higher operational leverage and profit volatility, lowering risk-adjusted profitability. Given the purpose of our model, the model's excellent fit to airline capacity, and the lack of publicly available data concerning

the aircraft supply line we chose not to include the structure for manufacturing capacity in our base model, but recognize it as an important issue for future research.

As discussed above, the wage sector could be elaborated to improve the fit of the model to the data for industry wages. Although the model fit to the data for wages exhibits an acceptable mean absolute error and low bias, it is the least accurate fit. During partial model testing we implemented several structures to attempt to improve the fit, including an experience chain to model the average tenure of the workforce. Since compensation rises as employees gain more years of experience, changes in the age distribution of the workforce would change average wages even if the wage at each pay step on the scale was constant. However, we found that average tenure was uncorrelated with the unexplained variation in average wages and therefore cut this structure from the model in the interests of parsimony. Determining what additional causal links, especially plausibly endogenous ones, would better model wages remains an opportunity for future research.

Similarly, we currently model the determinants of costs by representing jet fuel costs, average fuel efficiency, and wages explicitly because that level of aggregation is sufficient to fit total costs well. For example, since advertising has remained a very stable percentage of total operating expenses we did not need to include an explicit structure to represent the way advertising budgets are set. However, endogenous advertising decisions and related marketing efforts such as loyalty (frequent flyer) programs, might complete an interesting feedback with demand that we have omitted, especially in a model that portrayed individual airlines.

Conclusion

We develop an industry level model of airline profits that expands the boundary of prior models by including endogenous capacity, demand, prices, wages, costs, and profit. Methodologically our results document the first implementation of Markov chain Monte Carlo methods to estimate standard errors in a system dynamics model and highlight important limitations in current approaches to calibrating larger models.

Substantively, we find that price setting feedbacks play a surprisingly important role in determining industry profit stability. In particular, yield management strengthens a fast acting negative loop that damps the industry cycle in capacity, revenue and profit, but increases operational leverage and the volatility of profit. The net effect is to lower risk adjusted return compared to what would be possible using other pricing heuristics. Operational leverage is an important consideration in assessing risk-adjusted profit, but has not typically been modeled dynamically. While our model is not rich enough to address important dynamics that might arise from the implementation of different pricing rules at the inter-firm level, it provides a publicly-available platform for future researchers to explore these and other issues.

The methodological contributions of the work are less central, but there are several techniques that we fully document for the first time. In particular our paper is the

first to show how MCMC methods provide system dynamics research with a computationally efficient, general tool for determining the confidence intervals around parameter estimates. While the supply line adjustment for growth (OS2) and the data reporting structure (A1) that we use are not novel, the documentation we provide may be helpful for others who implement them.

Since deregulation more than thirty years ago airline profits have been close to zero on average and have experienced large amplitude cycles. Looking to the future, the increasingly commoditized nature of air travel, the rising costs of fuel, and the growing pressure to reduce industry greenhouse gas emissions suggest that airlines will likely face continued challenges to profitability. Our model is offered in the hope that it will be useful in the attempt to stabilize airline industry profits so that airlines can continue to provide a vital service in the global economy.

References

- Airbus 1998. *Global Market Forecast 1998-2017*, Airbus Industrie: Balagnac
- Aramburo S, Acevedo J, Morales Y. 2012. Laboratory experiments in the system dynamics field. *System Dynamics Review*, **28**, 94–106
- Belobaba P. 1987. Survey Paper, Airline Yield Management An Overview of Seat Inventory Control. *Transportation Science*, **21**(2), 63-73.
- Brown F. 2007. *Engineering system dynamics: a unified graph-centered approach (Vol. 8)*. CRC press.
- Carson R, Cenesizoglu T, Parker R. 2011. Forecasting (aggregate) demand for US commercial air travel. *International Journal of Forecasting*, **27**(3), 923–941.
- Croson R, Donohue K, Katok E, Sterman J. 2005. Order stability in supply chains: Coordination risk and the role of coordination stock. *Sloan School of Management Working Paper, Cambridge, MA*
- Daidsen P, Sterman J, Richardson G. 1990. A petroleum life cycle model for the United States with endogenous technology, exploration, recovery, and demand. *System Dynamics Review*, **6**(1), 66-93.
- Doganis R. 2002. *Flying Off Course: The Economics of International Airlines*. Routledge
- Ford A. 1997. System dynamics and the electric power industry. *System Dynamics Review*, **13**(1), 57-85.

- Forrester J. 1961. *Industrial dynamics*. Productivity Press, Cambridge, Mass.
- Forrester J. 1979. An alternative approach to economic policy: Macrobbehavior from microstructure. N. M. Kamrany, R. H. Day, eds. *Economic Issues of the Eighties*. Johns Hopkins University Press, Baltimore Maryland, 80-108.
- Forrester J. 2013. Economic theory for the new millennium (2003). *System Dynamics Review*. **29**(1): 26-41
- Forrester J, Mass N, Ryan C. 1976. The system dynamics national model: Understanding socio-economic behavior and policy alternatives. *Technological Forecasting and Social Change* **9**(1-2) 51-68.
- Gelman A, Carlin J, Stern H, Rubin D. 2003. *Bayesian data analysis*. Chapman & Hall/CRC.
- Ghaffarzadegan N, Tajrishi A. 2010. Economic Transition Management in a Commodity Market: The Case of the Iranian Cement Industry. *System Dynamics Review* **26**(2): 139-161
- Goncalves P. 2003. *Demand Bubbles and Phantom Orders in Supply Chains*. (Doctoral dissertation, MIT Sloan, 2003) Retrieved on 6/14/2012 from <http://web.mit.edu/~paulopg/www/PhD%20Thesis%20Goncalves.PDF>
- Groesser S, Schwaninger M. 2012. Contributions to model validation: hierarchy, process, and cessation. *System Dynamics Review*, **28**: 157–181
- Hadjis A. 2011. Bringing economy and robustness in parameter testing: a Taguchi methods-based approach to model validation. *System Dynamics Review*, **27**: 374–391
- Hansman R, Jiang H. 2005. *An Analysis of Profit Cycles in the Airline Industry*. Thesis (S.M.), Massachusetts Institute of Technology, Dept. of Aeronautics and Astronautics
- Homer J. 2012. Partial-model testing as a validation tool for system dynamics (1983). *System Dynamics Review*, **28**: 281–294
- Huefner R, Largay J. 2008. The role of accounting information in revenue management, *Business Horizons*, **51**(3): 245-255
- Kim J, Hyung Joon R, Jung G, Lee J. 2011. Convergence Monitoring of Markov Chains Generated for Inverse Tracking of Unknown Model Parameters in Atmospheric Dispersion. *Nuclear Science and Technology* (1) pp. 464-467.

Kimes S. 1989. Yield Management: A Tool for Capacity-Considered Service Firms. *Journal of Operations Management* Volume **8**(4): 348-363

Liehr M, Größler A, Klein M, Milling P. 2001. Cycles in the Sky: Understanding and Managing Business Cycles in the Airline Market. *System Dynamics Review* **17**(4): 311-332

Lyneis J. 2000. System Dynamics for Market Forecasting and Structural Analysis. *System Dynamics Review* **16**(1): 3-25

Makridakis S, Andersen A, Carbone R, Fildes R, Hibon M, Lewandowski R, Winkler R. 1982. The accuracy of extrapolation (time series) methods: results of a forecasting competition. *Journal of Forecasting*, **1**(2): 111-153.

Martinez-Moyano I. 2012. Documentation for model transparency. *System Dynamics Review*, **28**(2): 199-208

Mass N. 1975. *Economic cycles : an analysis of underlying causes*. Wright-Allen, Cambridge, Mass.

Meadows D. 1970. *Dynamics of Commodity Production Cycles*. Wright-Allen Press: Cambridge, MA.

Oum T, Waters W, Yong J. 1990. *A survey of recent estimates of price elasticities of demand for transport (Vol. 359)*. World Bank: Washington, DC.

Rahmandad H, and Sterman J. 2012. Reporting guidelines for simulation based-research in social sciences. *System Dynamics Review*, **28**(4): 396-411

Randers J, Göluke U. 2007. Forecasting Turning Points in Shipping Freight Rates: Lessons from 30 Years of Practical Effort. *System Dynamics Review*, **23**: 253–284.

Schafer A. 1998. The global demand for motorized mobility, *Transportation Research Part A: Policy and Practice*, **32**(6): 455-477

Sharpe W. 1994. The Sharpe Ratio. *The Journal of Portfolio Management* **21**(1): 49–58

Smith B. Laimkuhler J., and Darrow R. 1992 Yield Management at American Airlines. *Interfaces* **22**(1): 8-31

Stanovich K. 2011. *Rationality and the reflective mind*. Oxford University Press, USA.

Sterman J. 1984. Appropriate Summary Statistics for Evaluating The Historical Fit of System Dynamics Models. *Dynamica* **10**(2): 51-66

Sterman J. 1989. Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decision Making Experiment. *Management Science* **35**(3): 321-339.

Sterman J. 2000. *Business Dynamics: Systems Thinking and Modeling for a Complex World*. McGraw-Hill: Boston.

Sterman J, Fiddaman T, Franck T, Jones A, McCauley S, Rice P, Siegel L. 2012. Climate interactive: the C-ROADS climate policy model. *System Dynamics Review*, 28(3), 295-305.

U.S Department of Transportation, Federal Aviation Administration 2011. The Economic Impact of Civil Aviation on the U.S. Economy. Retrieved July 14, 2013, from http://www.faa.gov/air_traffic/publications/media/FAA_Economic_Impact_Rpt_2011.pdf

Weatherford L, Bodily S. 1992. A taxonomy and research overview of perishable-asset revenue management: yield management, overbooking, and pricing. *Operations Research*, **40**(5), 831-844.

Appendix A1: Data Reporting Macro

One of the challenges when fitting models to reported data is that data are typically reported at discrete intervals such as a month, quarter or year, while system dynamics models represent the underlying continuous dynamics of the system. Directly comparing reported data with the model's computed values can be problematic because the instantaneous value is not the same as the reported variable, which typically measures the average or accumulated value of a flow over some reporting period. For example, the instantaneous value of a corporation's revenue, net income, and other flow variables on a particular day will generally differ from the reported values on the income statement for the last quarter, since the reported values are the accumulated total for the period. The resulting errors can be substantial, and introduce the possibility of systematic bias in parameter estimation, particularly if there are trends in the data (for example, instantaneous sales and profit will be higher than the reported values for the last quarter when sales and profits are growing).

The model includes a structure, implemented as a Vensim macro, which replicates the data reporting process for a given¹⁰ data reporting period. The data reporting structure is shown in Figure A1.

¹⁰ The data reporting period must divide into 1 time unit with no remainder, e.g., 0.25 would represent four quarters/year.

The figure shows the example for the flow variable revenue (\$/year). The instantaneous, simulated revenue flows into the stock labeled “accumulated reported variable”. When the check reporting flag indicates that end of the reporting interval, say one quarter year, has been reached, the entire accumulated revenue over the quarter is removed from the stock. The value is then converted from the amount per reporting period to the amount per time unit used in the model by multiplying it by the reporting period (annualized when time is measured in years).

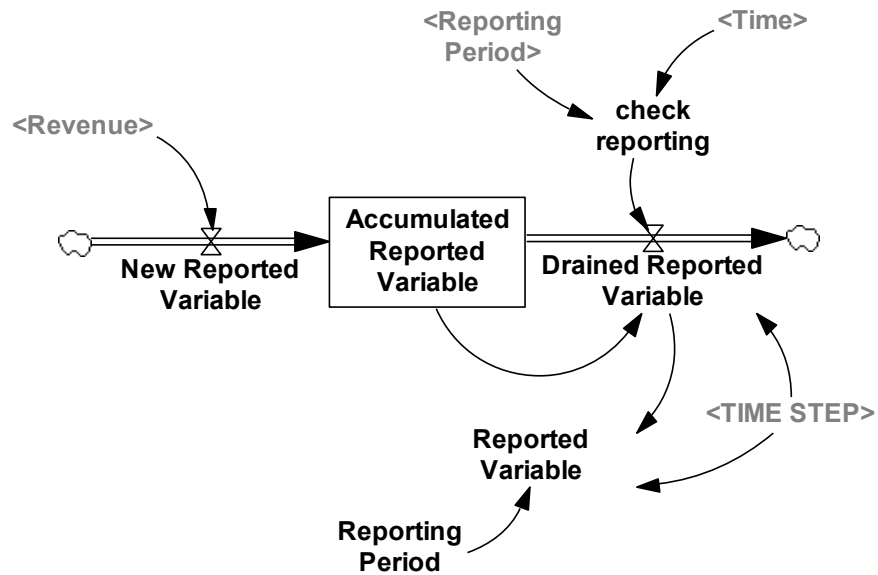


Figure A1: The data reporting structure implemented in the model

The Vensim macro that implements the data reporting structure is reproduced below:

```

:MACRO:Report Variable(Simulated Data, Reporting Period)
Report Variable = Drained Reported Variable*TIME STEP$/Reporting Period
~   Simulated Data
~   |
Drained Reported Variable=IF THEN ELSE(Check Reporting=0, Accumulated
Reported Variable/TIME STEP$, 0 )
~   Simulated Data
~   |
Check Reporting=MODULO(Time$, Reporting Period)
~   Reporting Period
~   |
Accumulated Reported Variable = INTEG(Simulated Data-Drained Reported
Variable,0)

```

```

~      Simulated Data*Reporting Period
~      |
:END OF MACRO:

```

Appendix A2: Markov chain Monte Carlo Standard Errors

Markov chain Monte Carlo (MCMC) standard errors are a recently implemented option for optimization in Vensim version 6. The algorithm assumes that the user has first found the globally optimum best-fitting parameters, and will return an error if it detects a set of parameters with a better payoff than its starting point.

During the process Vensim defines a region of parameter space that is “close” to the global optimum and selects new parameters following a random walk to efficiently find the range of values that jointly determine the confidence interval. The method requires two things. First, the payoff surface to be a likelihood or log likelihood; and second, the algorithm must run for long enough to adequately explore the space. The payoff function can be univariate or a weighted sum of the errors between simulated and actual data for multiple variables.

Defining a payoff as a log likelihood can be accomplished by setting the weight of the errors with respect to each dimension of the payoff to the reciprocal of the root mean square error of the simulated data (1/RMSE). (Gelman et al., 2003).

Doing so necessitates an iterative approach to optimization, because the mean square error between the data and the model isn’t known until the model has run. We defined a variable in the model that calculated 1/RMSE using the existing RMSE calculation in the summary statistics structure documented by Sterman (2000, Ch 21). Starting with arbitrary weights (we chose the inverse of the variance of the historical data), we ran a Powell optimization with multiple restarts and selected the set of parameters that fit best. We then replaced the weights on each data series with the realized inverse of the RMSE from that best fit. We then ran the Powell optimizer from this point without restarts, measured new weights given the resulting parameters, and started the optimizer again using the most recent parameters and weights. After approximately ten optimizations the process settled so that the changes in the parameters were much smaller than the threshold of three significant figures we set for precision. The estimated parameters at that point were declared the optima and the weights for running the MCMC algorithm were calculated using them.

Currently one must edit the .voc (vensim optimization control) files manually, using a text editor, to run the MCMC tests. Future versions of Vensim may implement MCMC within the user interface, but that work is not complete in version 6.1. Once the .voc file is edited manually, do not open it again from within the user interface or your manual edits will be lost. The following constitute the set of changes we made to the .voc files. There are many additional options inside of Vensim that can tune this process. The online documentation for version 6.1 (http://www.vensim.com/documentation/mcmc_sa.htm)

gives an extensive description and is a more complete resource than this appendix.

- :OPTIMIZER=Off
[Comment: turns off the Powell optimizer since you've already located the global optimum]
- :SENSITIVITY=Payoff MCMC=2
[Comment: Tells vensim to use the MCMC payoff, 2 is the boundary in the payoff space where vensim will define the 95% confidence interval and relies on your setting the payoff definition weights to 1/RMSE]
- :MCLIMIT= #number#
[Comment: This is optional, and if you set it to a negative number the optimizer will run until you turn it off.]
- :MCBURNIN= #number#
[Comment: This is also optional, and will discard a certain number of runs before attempting calculation of the MCMC distribution. The documentation recommends setting this to zero in most cases and then potentially increasing it if you run into strange results, but some burn-in is generally desirable.]

While the MCMC algorithm is running it will report potential scale reduction factors (PSRF) for each of your variables periodically. These diagnostics are used to determine when the MCMC has run for a sufficient time to be certain of the distribution. PSRF's are always greater than 1, and every variable should have a PSRF less than 1.2 before the process is terminated (Kim et al., 2011).

Two files will be output by Vensim once the payoff is complete. Runname_MCMC.tab reports the parameters and 95% confidence intervals. Runname_MCMC.dat has many additional diagnostics and a full report of the results.