Evaluating Risk for Mechanical Aviation Occurrences

Lawrence Berk

John A. Volpe National Transportation Systems Center,

55 Broadway, Cambridge, MA 02142

Abstract

Between January 1996 and March 2001 airline crews, following emergency procedures, prematurely ended approximately 33,000 flights due to problems with aircraft systems. Every event ended with safe return of the aircraft and no casualties. Adverse trends in these events, defined as *occurrences* by the FAA, may represent accident precursors. The U.S. DOT's Volpe Center maintains a real-time Intranet application, the Safety Performance Analysis System (SPAS), designed specifically for inspectors to assess risk, prioritize workloads and develop intervention strategies. FAA inspectors requested a SPAS tool highlighting both adverse and encouraging trends in occurrences specific to aircraft system, aircraft type and airline. A proposed technique based on contingency table analysis accomplishes this, employing surveillance data to isolate statistically significant differences among carriers in peer groups.

Keywords: Risk Analysis, Transportation, Aviation Safety

1. Introduction

Transportation Secretary Mineta has established the goal of "*…an 80% reduction in the U.S. commercial air carrier fatal accident rate* [measured per 100,000 departures] *by 2007.*" (DOT 2001, p. 40). In 2000, more than 2,700 air carriers possessed an active FAA certificate for passenger/cargo service under Title 14, Code of Federal Regulations (14 CFR) Part 121 (air transport) and/or Part 135 (commuter) (FAA 2001). FAA Inspectors typically specialize in one area: avionics, air carrier operations or aircraft maintenance. To meet goals of the Department and manage their workloads, FAA inspectors worldwide use the Safety Performance Analysis System (SPAS), developed for the FAA by the U.S. DOT's Volpe Center. SPAS provides daily access to every surveillance and inspection record and real-time statistical analyses on the surveillance and inspection data, usually integrating records from several sources. Statistically valid comparisons of data are of tremendous benefit to FAA inspectors, enabling them to develop effective intervention strategies. Of particular interest are assessments of risk airlines pose for the most common, yet least serious, of in-flight mishaps. Awareness of airline tendencies to experiences these problems – relative to industry peers – may lead to FAA interventions that reduce the risk of the more serious aviation mishaps, accidents and incidents.

An event in which the crew encounters an in-flight hazard, follows emergency procedures and prematurely ends a flight is an *occurrence* if the aircraft lands safely with very limited damage and no injuries among passengers and crew. (The FAA defines any event resulting in casualties or substantial damage to be either an *accident* or an *incident*, not an occurrence.) Since occurrences jeopardize the safety of passengers and crew, airlines and air traffic control facilities must notify the FAA of every event and each must be investigated by the FAA (FAA 2000). Occurrence investigation reports are not in the public domain, as occurrence investigations are considered surveillance activities (FAA 1997). Therefore, this article presents actual data in a disguised form.

FAA inspectors requested a trend analysis designed to identify both adverse and encouraging trends: the former to help develop intervention strategies, the latter to judge their impact and effectiveness and learn from peer carrier experience. A FAA Flight Standards Service report on aviation safety suggested an important area of research is the analysis of surveillance-based data as possible predictors of future aviation accidents (FAA 1997). Studies have found no evidence that individual carriers pose statistically significant differences in accident risk (Stouffer 1992; Barnett and Wang 2000). This article examines events that are mutually exclusive from accidents and incidents. Trends revealed in occurrences should not be interpreted as suggesting statistically significant differences in occurrence risk are correlated with, or are useful in, predicting accident risk.

This study is designed to support very specific surveillance activities carried out by inspectors. Occurrences by definition involve zero fatalities and so it is inappropriate to formulate risk assessments in terms of a probability of dying on a random flight (albeit useful when presenting safety risk from a public perspective (Barnett and Wang 2000)).

Since accidents and incidents are very rare, previous studies have evaluated safety between industry segments or geographic regions (Oster and Strong 1992; Barnett and Wang 2000). The pitfalls with these broad categorizations have been noted (Oster and Strong 1992, p. 27):

> *"[Overall rates] provide little guidance about where to focus effort to improve safety. A more promising approach begins with classifying accidents according to their cause and comparing the distribution of causes both over time and across segments of the industry."*

Accessible presentations of contingency table analyses permit inspectors to isolate very particular trends and identify opportunities for safety mitigation. The technique tests for deviations from an "equal-safety hypothesis" (Barnett and Wang 2000, p. 4) and whether these deviations can reasonably be ascribed to random chance. It is plausible to assume that the joint probability distribution of occurrences within a carrier peer group is the multinomial distribution. Given N independent trials across a peer group of air carriers, the chances carrier i experiences an occurrence or does not experiences an occurrence (p_i and 1- p_i, respectively) are estimated from the observed data, yet are assumed intrinsic and possibly unique. Therefore, marginal distributions of occurrences by carrier have mean Np_i, an assumption consistent with the widely held belief that differences in carriers' operations, policies and procedures greatly influence their respective chances of mishap. Contrast this assumption with a study of accidents, incidents and enforcement actions, in which the chance of mishap is hypothesized to be a constant (p) for an entire peer group. Samples consist of the total of M departures (each departure is treated as an independent trial) for a group of carriers over five- and ten-year periods along with n_i, the observed numbers of mishaps per carrier during the given period. Hypothesis tests comparing observed $\sum_i n_i = n$ group mishaps against a hypergeometric distribution of mean Mp establish whether individual carriers' experience events consistently with the assumption of a constant chance of mishap (GRA 1988). This premise leads to questionable conclusions regarding the safety risk posed by individual air carriers because they are all formulated from the premise that air carriers are 'all the same'. To assess and mitigate safety risk, FAA inspectors are particularly interested in *relative* comparisons of risk between air carriers, comparisons that explicitly recognize differences in carrier operations.

This article's main contributions to aviation safety data analysis are three-fold. First, it examines events not extensively covered in the literature. Every one of the referenced studies is limited to accidents or incidents, available to the general public. By contrast, this study examines events investigated as part of FAA inspector surveillance activities. A 1997 FAA Flight Standards Service report includes a comprehensive discussion of deliberations regarding potential publication of occurrence data (FAA 1997).

Since occurrences are by nature the most common type of event, more data is available from which to determine patterns and determine intervention. (If it is found that occurrence precursors are also precursors to more serious accidents and incidents, there is the potential for tremendous benefit from FAA intervention to the flying public.) By analyzing occurrence data, another contribution of this study is that it demonstrates persistent patterns can be isolated by carrier at greater levels of detail and over shorter time periods than reported by accident studies. Aviation safety analyses generally aggregate event data by examining longer time periods or limit analyses to carrier peer groups to compensate for the rarity of accidents and incidents. Indeed, every one of the studies cited in this article infer safety risk from event rates spanning well in excess of a calendar year. Mishap rates by carrier may span as much as a decade of flight (Barnett and Wang 2000) or annual event rates are derived strictly by carrier groups (GRA 1988).

The choice of data and method suggest the potential to isolate patterns in event risk that are both persistent and specific enough to assist FAA inspectors in developing proactive strategies. This speaks to the article's third contribution – a unique presentation of risk analysis from the perspective of government oversight and monitoring of aviation safety.

This article is organized around five sections. The next section discusses source data and its collection. The third section describes key assumptions and validation techniques. The fourth section describes the contingency table approach and defines a supplemental risk measure. Computational results and resulting inferences are summarized in the fifth section. The paper concludes with final comments and suggestions for further research.

2. Data

2.1 Occurrence Records

Carriers and air traffic control facilities must notify the FAA of every mishap, however serious (FAA 2000). Consequently, no bias is introduced due to voluntary reporting, in contrast with previous aviation safety studies examining less serious events (Villareal 1988; Stouffer 1992). Only events in which investigations conclude the primary cause was a mechanical failure are included in this analysis, which excludes occurrences caused by medical emergencies or events strictly resulting from pilot error or air traffic control. Mechanical events almost always involve one aircraft; damage due to collision is not associated with failure of a mechanical component. Since commercial air carriers have operational responsibility for their entire fleet, regardless of leasing or contract maintenance agreements, it is natural to associate each mechanical occurrence with the carrier. Such an association is less plausible given events due to medical emergencies or pilot deviations. By nature, these often involve factors beyond carriers' operational control.

Each occurrence investigation report is very detailed, including information on: time, location, operator, aircraft, flight personnel, nature of mechanical fault, FAA inspector(s) and a descriptive narrative of the occurrences and subsequent actions. The precise nature of a mechanical fault is described in reports by standard 4-digit Air Transport Association (ATA) codes categorizing aircraft components. The first two digits, the *ATA Chapter*, refer to aggregate systems, e.g., "Flight Controls", "Lights", and "Powerplant". FAA inspectors specified that aggregation of occurrences by ATA Chapter is appropriate for comparisons and trend analyses and will be sufficient for developing mitigation strategies.

This technique is intended for a FAA software system which must analyze and present data in real-time. Occurrence records that omit mechanical fault details or aircraft type or mechanical fault are simply discarded. Preliminary analyses of the source data strongly suggest no algorithm harvesting details from report narratives would be sufficiently error-free; the time and effort required would also be prohibitive.

An inspector team may conduct an occurrence investigation. Either expertise in multiple specialties is required or – as is very often the case – an occurrence takes place outside the jurisdiction of the responsible inspector(s). FAA policies require an investigation report from both an inspector 'on scene' and from the inspector(s) principally responsible for oversight. Every occurrence investigation report must be entered into FAA data systems, which simply count every report as an unique event.

Because this practice has the potential to dramatically overstate carrier risk, source reports were first manually reviewed, and counts revised downward so that only one report is counted per mishap. Comparisons of the report narrative, important dates, and various database fields (e.g., aircraft registration number, airline flight number, aircraft model, origin/destination airport) were often sufficient to determine conclusively whether multiple reports referred to one event. When there was any doubt, this study assumed reports refer to the same event, reducing carrier occurrence counts, implying conservative estimates of risk.

2.2 Utilization Records

Each of the referenced aviation safety studies includes a discussion of the merits and pitfalls of common risk exposure measures: passenger-miles, flight-hours and number of departures. The typical argument against passenger-miles as an exposure measure is that it discriminates against carriers flying shorter trips or fewer passengers (Stouffer 1992; FAA 1997). There are similar concerns with hour-based metrics (e.g., flight hours), which may discriminate against carriers employing faster aircraft or do not account for differences in customer base (Villareal 1988).

This study uses 'flight hours' (or 'utilization) as a proxy for carrier risk exposure for two reasons. First, exposures of mechanical components to risk of failure are best expressed as estimates of time in service. The chance of mechanical fault is assumed proportional to a component's service time. The more it is used, the greater the chance of failure; the less it is used, the lower the chance of failure. With but few exceptions (e.g., ATA Chapter 32 – Landing Gear), equating service time with the number of takeoffs and landings (cycles) is inadequate.

Second, the industry's recent experience with hazardous aviation events suggests they have occurred in comparable proportions between the early, cruise and late stages of flight. A common argument in accident study literature, favoring cycles and against utilization, claims a disproportionate number of accidents occur during the very early or very late stages of a flight (Stouffer 1992; FAA 1997). Short of reading every occurrence investigation narrative, it is not possible to prove or disprove whether this claim is also true of occurrences; 'phase of flight' is not a database field in the FAA occurrence investigation database. However, an analysis of events that includes incidents – less serious than accidents, yet involving injury and substantial aircraft damage – supports the earlier claim and lends credence to the use of utilization as a risk exposure proxy. (The Appendix presents these findings in greater detail.)

The U.S. DOT (FAA 2001) collects monthly utilization (flight hour) statistics by air carrier and aircraft type. In this study, hours are first aggregated by aircraft Make-Model (e.g., B-737) and are not retained by Make-Model-Series (i.e., B-737-205, B-737-400). Aircraft series designations are often, but not always, chronological; aggregated statistics conceivably equate aircraft of possibly disparate ages. A plausible hypothesis is that aircraft age contributes to occurrence risks associated with mechanical failure. Unfortunately, in the source data, aircraft series are either a) inconsistently recorded due to carrier reporting practices (i.e., some carriers aggregate all B-737 hours, others distinguish by B-737 series) or b) prone to reporting errors (e.g., B-737-317 is a valid series, but B-737-371 is not). Furthermore, only age of the airframe can be approximated from the series, not necessarily ages of every aircraft component. Aircraft are continually upgraded and maintained with new replacement parts and engines that are often interchangeable between aircraft series. The source data also does not distinguish between flight hours logged on domestic and international flights. However, this distinction is irrelevant since every carrier required to report to DOT possesses an FAA certificate, and is therefore subject to FAA oversight.

Monthly utilization (flight hour) reports by aircraft model, even among airlines required by 14 CFR to file, can be incomplete in the source data. In the great majority of cases, only one or two data points were missing for active carriers. Except for one active carrier, fewer than six month's of reported data in a year always implied the carrier utilized that type of aircraft only part of the year (i.e., the carrier introduced or retired aircraft).

Time series plots of a carrier's flight hours by aircraft model strongly suggested that simple regression models with time as the only independent variable would adequately capture common growth and seasonal trends in utilization. Consequently, simple regression models offer reasonable estimates for missing data that are superior to the sample average. (Explaining flight hour variability by carrier and aircraft type was not the intent.) Polynomials in time as an ordinal variable (an index variable for the month) are fit to the available data. Denote the number of months that are (or would be) spanned by a complete set of carrier A monthly flight hours by n_A. Define the index set $T_A = \{1,2,\ldots,n_A\}$. Let index set $M_A \subset T_A$ denote months corresponding to missing values; let index set $K_A = T_A \setminus M_A$ denote months corresponding to reported values. A vector, ***b***, of least-squares coefficient estimates for a degree-D polynomial in ordinal time variable $k \in K_A$ is derived from the regression equation:

$$\begin{bmatrix} h_A(\underline{k}) \\ \vdots \\ h_A(k) \\ \vdots \\ h_A(\bar{k}) \end{bmatrix} = \begin{bmatrix} 1 & \underline{k} & (\underline{k})^2 & \cdots & (\underline{k})^D \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & k & k^2 & \cdots & k^D \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \bar{k} & (\bar{k})^2 & \cdots & (\bar{k})^D \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_D \end{bmatrix} + \begin{bmatrix} e_{\underline{k}} \\ \vdots \\ e_k \\ \vdots \\ e_{\bar{k}} \end{bmatrix}$$

where $h_A(k)$ denotes the flight hours recorded for carrier A in month $k \in K_A$, $\underline{k} \equiv \min_{k \in K_A} k, \bar{k} \equiv \max_{k \in K_A} k$ and error terms are independent, identically distributed $e_k \sim N(0,\sigma^2)$.

Values of n_A and D were judiciously set by inspection of the data. Usually $n_A = 12$ (i.e., the set spanned one calendar year) when source data possessed only one or two missing values. If data were missing early (late) in the year, the set of data points is augmented *a priori* with data points late (early) in previous (subsequent) years so that $n_A > 12$. This practice enabled polynomial models to capture evident year-to-year trends and cycles in a carrier's flight hours. In the great majority of cases, $D \in \{3,4,5\}$.

These data collection rules and model forms generate superior estimates vis a vis sample averages and are very simple to implement and maintain in real-time computer software. Missing value estimates can be readily obtained in the intended real-time software either through a regression routine that chooses from multiple polynomial models of varying degree (say, with aid of an adjusted R^2

statistic), or merely fitting the least-squares fifth degree polynomial (an example of the necessary tradeoffs between complete rigor and software requirements).

3. Approach

Define a *population* $P_{m,t}$ as the complete set of commercial air carriers that operate at least one aircraft of type m during period t. As time passes $P_{m,t}$ is generally dynamic as carriers retire or introduce type m aircraft. FAA inspectors provided practical criteria for defining t and subdividing large populations $P_{m,t}$ into manageable *peer groups* $g_{m,t} \subseteq P_{m,t}$. Members of $g_{m,t}$ possess operational similarities and must conform to similar sections of 14 CFR.

Theoretically, it may be possible to collect the precise start and end times of an occurrence (e.g., from flight data recorders). Times to such levels of precision rarely, if ever, find their way into occurrence investigation reports. Typically, an occurrence 'start time' coincides with the approximate time a crew notified air traffic control of a problem. An occurrence 'end time' coincides with the approximate time the aircraft touched down. Elapsed times recorded in investigation narratives were reviewed from dozens of reports. In every narrative, occurrences lasted well under one hour. This is certainly not surprising, since occurrences are safety hazards necessitating emergency actions on the part of the crew. We therefore assume that an occurrence spans at most one flight hour. Consequently, the study defines an *observation* as a single, randomly selected, flight hour. This implies that counts of carriers' observed occurrences equal the same numbers of observed flight hours. If a carrier experienced more than one occurrence on the same day due to the same fault (virtually nonexistent circumstances that always involved multiple aircraft), each event is assumed to run its course during non-overlapping hours.

Construct a 2x| $g_{m,t}$ | contingency table as follows. Let A denote the set of ATA Chapters. For a specified aircraft type m, time period t and ATA Chapter $f \in A$, elements of the first row count $o_{c,f}$, the number of flight hours in which carrier $c \in g_{m,t}$ experienced an occurrence caused by an $f \in A$ system fault. Elements of the second row tabulate the remainder, $r_{c,f} = H_{c,\bullet} - o_{c,f}$, where $H_{c,\bullet}$ is total flight hours (reported or estimated), independent of fault $f \in A$ for carrier $c \in g_{m,t}$ in aircraft type m during period t.

Contingency table cells form a partition of mutually exclusive outcomes (an occurrence or uneventful flight hour) across carriers $c \in g_{m,t}$, given $f \in A$.

By assumption, cell probabilities follow a multinomial distribution. For each $c \in g_{m,t}$ and $f \in A$, the joint probability distribution of cell outcomes is

As a group, operators of Type I aircraft experienced comparable annual landing gear occurrences between 1997-2000 (between 109 and 116), significantly more than the 1996 group total (88). Carrier F increased utilization year-over-year, operating the aircraft at levels far in excess of any of its peers each year. Through 1999, carrier F had not presented significantly high risk for landing gear mishaps. Yet, as carrier F reached peak utilization in 2000, risk for mishaps increased, becoming statistically significant through the first quarter of 2001. By contrast, mishap risk for carriers A, B and C, all with year-over-year increases in utilization and totals exceeding 500,000 hours, were not significantly high as they reach peak utilization. Carrier F is on pace to experience more than fifty mishaps by January 2002 – 25% more than during 2000 – against a modest projected increase in annual utilization. Note that, relative to carrier F's 1996 *Risk* (1.4), *Risk* was lower (1.2) in 1998 and 2000 despite more mishaps. Both can be explained by the dramatic increase in group occurrences. In 1998, carrier F's fraction of group utilization (29.1%) was relatively unchanged from 1996 (31.1%). Yet, because the group experienced 110 occurrences in 1998, and only 88 in 1996 *Risk* numerators decreased; proportionally, carrier F's contributions to group mishaps declined over time.

Carrier L introduced Type I aircraft in 1999 and posed significantly higher-than-expected risk. It is plausible that inexperience of the carrier's pilots and ground crews contributed to the high number of occurrences that first year and relative improvement since. (Yet, during 2000, carrier L's seven events were still indicative of high risk.) Despite the benign risk level during the first quarter of 2001, carrier L may deserve continued scrutiny. The dramatic increase in Type I aircraft utilization was accompanied by more events. Projecting forward, carrier L could experience more events in 2001 than in 2000 as annual 2001 utilization may exceed 2000 levels.

Carrier K took a significant hiatus from operating Type I aircraft in 1998. As with carrier L, carrier K presented dramatically high risk when service resumed in 1999; more than seven times the group average! Since then carrier K curtailed flights of Type I aircraft; utilization dropped to a fraction of what it was, and no events occurred from January 2000 through March 2001. Carriers E and G (both regional) each experienced nine landing gear events in 1999 and posed significant risk, considering their respective annual Type I utilization figures are among the lowest of the group. (For example, both carriers E and G experienced more 1999 events their either trunk carriers C or A; the latter flew the Type I aircraft six to eight times as often.) Since 1999, both regional carriers experienced fewer events attendant with modest increases in Type I aircraft utilization, perhaps due to FAA intervention.

Aviation safety inspectors not only want to highlight high-risk carriers for targeted intervention, but seek to learn from carriers with better safety records for similar events. Table 1 suggests trunk carrier J, with high Type I aircraft utilization, posed lower–than–expected risk four years in a row. Annual carrier J

Similarly, carrier M experienced fewer than the expected number of between eight and eighteen occurrences per year. In 1996 and 1998, years when carrier M did not distinguish itself, the carrier reported multiple cabin pressurization problems or warnings of excessively high temperatures in the tail cone section (all resulting in immediate turnbacks).

4.3 Jetliner Type III

Table 3 presents a risk assessment for occurrences due to mechanical failure of Type III aircraft flight control systems (ATA Chapter 27). Typical mishaps describe control problems with leading edge slats, trailing edge flaps, wing spoilers (which aid ailerons on later-generation jetliners) or the rudder assembly. Missing values for Carriers B and C correspond to years when both airlines had not yet introduced the Type III aircraft.

Since the Type III aircraft is a larger, wide-body aircraft – generally relegated to long-haul domestic and international routes – the peer group is relatively small, aircraft utilization is relatively low, and mishaps are particularly uncommon. Consequently, every year, carrier risks are generally consistent with the expectation that mishaps occur in proportion to utilization. Carrier A stands out; recent risk for mishaps due to flight control problems was rather high, yet fleet utilization remained stable. Through the first quarter of 2001, mishap risk was insignificant. But, investigation records show that by September, carrier A experienced 14 additional occurrences (for a total of 19). By that time, carrier A experienced more mishaps with its Type III fleet than carrier D, yet carrier D had overtaken carrier A in Type III utilization. Carrier E increased utilization four-fold between 1996 and 2000, yet never posed significant mishap risk. Neither carriers B nor C experienced a mishap due to flight controls since their introduction of the Type III aircraft.

5. Concluding Remarks

Studies of accident and incident investigations have conclusively demonstrated that safety risk between individual carriers is indistinguishable. Experience with accidents is not indicative of persistent risk and cannot predict future safety. This study offers a complementary assessment of risks for occurrences, the least serious and most common of aviation mishaps. Analyses can be performed in greater detail and over shorter time horizons than practical with accidents and incidents.

Persistent patterns in annual occurrence rates for a particular fault, aircraft and carrier can be revealed. This finding is particularly important, since it suggests inspectors may use this technique to develop insights about industry safety, prioritize workloads and develop intervention strategies. If historical risk patterns persist up to the current time, there may be reason to suspect this risk will continue. However, prediction is beyond the scope of this analysis and the tool is not intended to serve in that capacity.

This study also considered occurrence risk between regional carriers operating a very popular turboprop Make-Model. However, risk inferences are biased toward carriers that report utilization. (Pursuant to 14 CFR, only aircraft with a maximum operating capacity of sixty seats or 18,000 pounds must have their utilization figures reported (OFR 2001).) This particular aircraft does not meet these standards. Error is compounded by a greater reliance on regression estimates (carriers report sporadically) to derive annual figures among those carriers that do report utilization. Consequently, in the absence of more complete utilization data (or an alternative), risk levels for small regional carriers are likely to be significantly misstated. Finally, this technique must be modified if is to be applied to General Aviation data. The FAA employs the General Aviation and Air Taxi Activity (GAATA) Survey to estimate utilization, strictly for generic aircraft (e.g., Piston Single-Engine, Turboprop). Furthermore, the number of commercial, municipal and recreational holders of FAA certificates renders analyses by individual operator impractical.

There is no evidence that persistent patterns in carrier *occurrence* rates are positively correlated with or can predict carrier *accident* rates. Indeed, it has been shown that rates for less-serious accidents and incidents can be negatively correlated with those for fatal events when examined by carrier (Barnett and Wang 2000). The assumption that patterns in occurrences due to a particular mechanical fault can predict accident risk associated with the same fault is too simplistic. (Mishaps are evaluated against one causal variable.) This method also cannot *rank* carriers by occurrence risk. Each risk measure $R_{c,f}$ is the ratio of actual to expected occurrences for carrier c due to fault $f \in A$. It is readily apparent from the assessments that a carrier posing expected risk (possessing an insignificant p-value) may possess a higher value of $R_{c,f}$ than another carrier with significant risk.

Finally, the examples examine carrier risk *given* $O_{\bullet,f}$, the number of events experienced by the group. No inferences were drawn about patterns in group risk, and individual risk is not assessed in complete context. However, by applying the technique to 2xY contingency tables, one may assess group risk patterns over Y periods. Marginal totals count group flight hours by period (by column) and outcome (by row). For example, an analysis of group air conditioning system (ATA Chapter 21) occurrences for Type II aircraft confirms what is apparent from Table 2. From 1996 into 2001 (six periods), the group experienced an alarming increase in occurrences, particularly since 1999; meanwhile group aircraft utilization declined. The p-value for the group Chi-Squared statistic with five degrees of freedom is significant at the $\alpha = 0.01$ level; so are Z-scores for 2000 and first-quarter 2001 (when the group experienced 125 and 36 mishaps, respectively). This analysis makes the point more strongly that, by comparison, 1996 was the safest year for Type II operators (88 mishaps) and that group safety has worsened every since.

References

Barnett, A., Wang, A. (2000), “Passenger-mortality Risk Estimates Provide Perspectives About Airline Safety,” *Flight Safety Digest* 19, 4, 1-10.

Everitt, B.S. (1977), “The Analysis of Contingency Tables,” London: Chapman and Hall.

Federal Aviation Administration (1997), “Aviation Safety Accessibility Study,” Washington, DC: Federal Aviation Administration, (http://www/asy.faa.gov/safety_analysis/aviationstudy.htm)

———(2000), “Aircraft Accident and Incident Notification, Investigation and Reporting,” Washington, DC: U.S. Department of Transportation.

———(2001), National Vitals Information Subsystem. Available: http://av-info.faa.gov.

Gellman Research Associates, Inc. (1988), “Analysis of Safety Differences among Air Carriers,” Report to Volpe National Transportation Systems Center.

Office of the Federal Register (2001), “Code of Federal Regulations,” Washington, DC: National Archives and Records Administration, (http://www.access.gpo.gov).

Oster, C. V., Strong, J.S. (1992), “The Worldwide Aviation Safety Record,” *Logistics and Transportation Review*, 28, 1, 23-48.

Stouffer, V. (1992), “Commercial Aviation Safety and Risk,” *Transportation Research Record*, 1332, 40-47.

U.S. Department of Transportation (2001), “Performance Report (FY 2000) Performance Plan (FY 2002),” Washington, DC: U.S. Department of Transportation, (http://ostpxweb.dot.gov/budget/PerfPlan02.pdf).

Villareal, C. (1988), “Uses and Misuses of Risk Metrics in Air Transportation,” *Transportation Research Record*, 1161, 31-42.

Wachtel, A., Lewiston, C. (1994), “Risk Factors for Bicycle-Motor Vehicle Collisions at Intersections,” *ITE Journal*, 64, 9, 30-35.

Table 1. Risk Assessment of Landing Gear Occurrences with Type I Aircraft

More than 500K flight hours						Fewer than 500K flight hours					
Carr	*Year*	*Occ*	*Hrs*	*Risk*	*p*	*Carr*	*Year*	*Occ*	*Hrs*	*Risk*	*p*
F	1996	36	7.8	1.4	0.01	D	1996	1	0.3	1.2	0.87
	1997	31	8.5	0.9	0.72		1997	2	0.3	1.5	0.56
	1998	38	9.0	1.2	0.22		1998	1	0.4	0.6	0.65
	1999	33	9.9	1.0	0.87		1999	3	-	-	-
	2000	41	11.1	1.2	0.11		2000	2	0.7	0.9	0.90
	1Q 2001	13	3.0	1.5	0.00		1Q 2001	1	0.2	1.8	0.56
J	1996	11	5.7	0.6	0.05	G	1996	0	0.3	0.0	0.29
	1997	14	5.9	0.6	0.05		1997	2	0.3	1.5	0.56
	1998	12	6.9	0.5	0.01		1998	3	0.3	2.4	0.11
	1999	12	7.4	0.5	0.00		1999	9	0.3	7.7	0.00
	2000	13	6.1	0.7	0.20		2000	3	0.5	2.2	0.15
	1Q 2001	3	1.3	0.8	0.51		1Q 2001	0	0.1	0.0	0.58
H	1996	16	6.3	0.8	0.26	L	1996	-	-	-	-
	1997	28	6.3	1.2	0.40		1997	-	-	-	-
	1998	14	5.6	0.7	0.13		1998	-	-	-	-
	1999	6	5.3	0.3	0.00		1999	3	0.3	2.9	0.05
	2000	12	5.1	0.8	0.37		2000	6	1.1	1.8	0.14
	1Q 2001	2	1.3	0.5	0.10		1Q 2001	2	0.4	1.7	0.46
B	1996	10	4.0	0.8	0.38	E	1996	0	0.2	0.0	0.41
	1997	15	4.2	0.9	0.76		1997	1	0.3	1.0	0.98
	1998	21	4.7	1.2	0.28		1998	0	0.2	0.0	0.36
	1999	27	5.9	1.3	0.12		1999	8	0.3	8.2	0.00
	2000	15	6.9	0.7	0.21		2000	3	0.3	3.2	0.03
	1Q 2001	3	1.8	0.6	0.18		1Q 2001	0	0.1	0.0	0.60
C	1996	6	1.1	1.6	0.21	K	1996	0	0.3	0.0	0.36
	1997	15	2.0	1.9	0.01		1997	0	0.3	0.0	0.31
	1998	14	2.2	1.8	0.03		1998	-	-	-	-
	1999	6	2.6	0.7	0.33		1999	6	0.2	8.2	0.00
	2000	8	3.0	0.9	0.75		2000	0	0.1	0.0	0.62
	1Q 2001	4	0.8	1.8	0.22		1Q 2001	0	0.02	0.0	0.81
A	1996	8	1.1	2.2	0.02						
	1997	5	1.2	1.0	0.92						
	1998	7	1.5	1.3	0.45						
	1999	6	1.7	1.0	0.92						
	2000	6	1.9	1.1	0.90						
	1Q 2001	0	0.5	0.0	0.19						

Table 2. Risk Assessment of Air Conditioning System Occurrences with Type II Aircraft

More than 1M flight hours						Fewer than 1M flight hours					
Carr	*Year*	*Occ*	*Hrs*	*Risk*	*p*	*Carr*	*Year*	*Occ*	*Hrs*	*Risk*	*p*
A	1996	36	8.2	1.9	0.00	C	1996	0	1.5	0.0	0.05
	1997	40	8.3	1.9	0.00		1997	5	1.5	1.3	0.54
	1998	32	8.3	1.4	0.02		1998	2	1.5	0.5	0.30
	1999	39	8.0	1.7	0.00		1999	3	1.4	0.8	0.61
	2000	49	5.8	1.6	0.00		2000	7	1.2	1.2	0.67
	1Q 2001	12	2.1	1.2	0.40		1Q 2001	2	0.4	1.1	0.86
K	1996	10	4.2	1.0	0.96	R	1996	4	-	-	-
	1997	12	4.2	1.1	0.67		1997	1	0.7	0.6	0.59
	1998	7	4.1	0.6	0.18		1998	2	0.7	1.0	0.95
	1999	10	4.1	0.9	0.59		1999	6	1.0	2.0	0.08
	2000	20	4.1	1.0	0.98		2000	14	0.9	2.9	0.00
	1Q 2001	8	1.4	1.3	0.35		1Q 2001	4	0.3	2.8	0.03
M	1996	7	4.2	0.7	0.31	B	1996	0	0.6	0.0	0.24
	1997	1	3.2	0.1	0.01		1997	0	0.6	0.0	0.20
	1998	7	3.6	0.7	0.34		1998	0	0.7	0.0	0.17
	1999	2	3.7	0.2	0.00		1999	1	0.7	0.5	0.48
	2000	5	3.5	0.3	0.00		2000	0	0.7	0.0	0.06
	1Q 2001	0	1.1	0.0	0.01		1Q 2001	0	0.2	0.0	0.30
E	1996	0	3.4	0.0	0.00	L	1996	0	0.8	0.0	0.17
	1997	0	3.4	0.0	0.00		1997	1	0.8	0.5	0.49
	1998	1	3.4	0.1	0.00		1998	2	0.7	1.1	0.93
	1999	3	3.3	0.3	0.02		1999	0	0.4	0.0	0.27
	2000	2	3.0	0.1	0.00		2000	-	-	-	-
	1Q 2001	2	0.8	0.5	0.33		1Q 2001	-	-	-	-
D	1996	11	2.9	1.6	0.09	G	1996	1	0.2	2.3	0.39
	1997	5	2.9	0.7	0.33		1997	0	0.2	0.0	0.44
	1998	9	2.6	1.2	0.49		1998	2	0.2	3.0	0.10
	1999	2	2.3	0.3	0.06		1999	2	0.5	1.4	0.66
	2000	7	1.8	0.7	0.40		2000	2	0.6	0.6	0.48
	1Q 2001	1	0.6	0.4	0.27		1Q 2001	5	0.2	4.5	0.00
N	1996	2	2.1	0.4	0.16	H	1996	0	0.2	0.0	0.47
	1997	7	2.3	1.2	0.60		1997	1	0.2	2.0	0.49
	1998	13	2.0	2.4	0.00		1998	0	0.2	0.0	0.46
	1999	12	1.8	2.3	0.00		1999	1	0.2	1.6	0.65
	2000	16	1.5	2.0	0.00		2000	2	0.2	1.6	0.48
	1Q 2001	2	0.4	1.2	0.81		1Q 2001	0	0.1	0.0	0.57
						J	1996	-	-	-	-
							1997	0	0.1	0.0	0.52
							1998	-	-	-	-
							1999	0	0.2	0.0	0.43
							2000	1	0.6	0.2	0.10
							1Q 2001	0	0.1	0.0	0.30

Table 3. Risk Assessment of Flight Control Occurrences with Type III Aircraft

More than 300K flight hours						Fewer than 300K flight hours					
Carr	*Year*	*Occ*	*Hrs*	*Risk*	*p*	*Carr*	*Year*	*Occ*	*Hrs*	*Risk*	*p*
D	1996	3	2.3	0.5	0.15	H	1996	1	0.5	0.8	0.78
	1997	7	2.7	0.9	0.65		1997	2	0.5	1.3	0.69
	1998	14	3.2	1.3	0.18		1998	0	0.5	0.0	0.17
	1999	13	3.8	1.1	0.73		1999	0	0.6	0.0	0.16
	2000	8	3.7	0.8	0.38		2000	0	0.6	0.0	0.19
	1Q 2001	8	1.1	1.5	0.33		1Q 2001	1	0.1	1.6	0.71
A	1996	12	3.1	1.5	0.08	E	1996	0	0.2	0.0	0.50
	1997	12	3.1	1.3	0.26		1997	1	0.4	0.9	0.90
	1998	13	3.1	1.3	0.34		1998	0	0.5	0.0	0.18
	1999	15	3.2	1.5	0.05		1999	0	0.7	0.0	0.12
	2000	14	3.3	1.5	0.06		2000	3	0.8	1.3	0.60
	1Q 2001	5	0.8	1.3	0.47		1Q 2001	0	0.2	0.0	0.26
G	1996	4	1.8	0.8	0.71	B	1996	-	-	-	-
	1997	5	1.8	0.9	0.90		1997	-	-	-	-
	1998	4	1.8	0.7	0.39		1998	0	0.01	0.0	0.86
	1999	4	1.9	0.7	0.37		1999	0	0.1	0.0	0.52
	2000	4	2.0	0.7	0.43		2000	0	0.3	0.0	0.39
	1Q 2001	1	0.5	0.4	0.22		1Q 2001	0	0.1	0.0	0.47
F	1996	2	0.6	1.3	0.67	C	1996	-	-	-	-
	1997	0	0.6	0.0	0.17		1997	-	-	-	-
	1998	1	0.7	0.4	0.37		1998	-	-	-	-
	1999	2	0.7	0.9	0.85		1999	-	-	-	-
	2000	3	0.7	1.4	0.52		2000	0	0.04	0.0	0.74
	1Q 2001	0	0.2	0.0	0.33		1Q 2001	0	0.1	0.0	0.47

APPENDIX: Proportions of Aviation Events by Phase of Flight

Investigation records from the FAA's Accident/Incident Database System (A/IDS) are compiled by reported phase of flight (Table A.1). A significant number of A/IDS investigation records have 'Unknown' phase of flight or are omitted entirely. These entries are most likely a consequence of draft reports having been submitted to A/IDS early on in the investigation (very soon after inquiries on-scene, yet well in advance of rigorous examination of circumstances and determination of cause). As with all definitions of 'takeoff' and 'landing', this classification is subject to interpretation. The data of Table A.1 are classified into four broad categories: events during 'Takeoff/Landing', 'Enroute', 'Airport Operations' and 'Unknown' (Table A.2). The category 'Takeoff/Landing' comprises all phases of flight with an asterisk (*) in the second column of Table A.1. This category represents the collection of phases that are likely to have occurred very early or very late in a scheduled flight. The category 'Airport Operations' comprises all events for phases of flight *Ground Taxi, Other Airplane* and *Other Ground Ops*. The category 'Unknown' comprises all events with an *Unknown* or *Unreported* phase of flight. Finally, events in category 'Enroute' comprise all remaining phases.

Clearly, recent events have not occurred disproportionately during the 'Takeoff/Landing' category; certainly not the oft-quoted "70 percent" statistic in accident studies (Stouffer 1992, p. 41; FAA 1997, p. 5). Indeed, among events for which phase of flight was known, very comparable proportions occurred in the 'Enroute' phase (cruising segments). We can only assert that 70% of events occur very early or late in a flight if we include in our 'Takeoff/Landing' definition a) all events involving ground support vehicles or repositioned aircraft and b) all events with no recorded phase (based on an assumption the empirical data does not support).

Table A.1. Recent Accidents and Incidents by Phase of Flight

A/IDS Phase of Flight		No. Events 2000	2001
UNREPORTED		119	-
STARTING ENGINES	*	5	3
IDLING ENGINES	*	4	1
ENGINE RUN-UP	*	2	1
POWER ON DESCENT (ROTORCRAFT)	*	1	-
IN TRAFFIC PATTERN-CIRCLING		1	-
FINAL APPROACH	*	48	25
INITIAL APPROACH-IFR	*	11	8
FINAL APPROACH-IFR	*	2	1
MISSED APPROACH-IFR	*	-	1
CLIMB TO CRUISE		89	56
NORMAL CRUISE		190	113
FCD/PREC LDG FROM CRUISE		7	-
LOW LEVEL OPERATIONS		1	-
DESCENT	*	37	19
AERIAL TAXI TO TKOF/RTRCRFT	*	2	-
GROUND TAXI, OTHER AIRPLANE		102	41
OTHER GROUND OPS		31	15
HOVERING		2	-
LEVEL OFF TOUCHDOWN	*	50	41
ROLL-OUT (FIXED WING)	*	50	28
PWR ON VERTICAL LDNG/RTRCRFT	*	2	-
PWR OFF VERT LAND/AUTOROTATE	*	1	-
PRW ON RUN LANDING (ROTORCRAFT)	*	2	-
TAKEOFF GROUND ROLL	*	41	37
TO INIT CLIMB(1ST PWR REDUCT)	*	68	27
TO VERTICAL (HELICOPTER ONLY)	*	2	1
TO RUNNING(HELICOPTER/VTOL AC)	*	1	-
TO ABORTED (FIXED WING)	*	12	7
FORCED/PRECAUTIONARY LANDING		1	1
UNKNOWN		2	1

NOTE: Phases of flight marked with an asterisk (*) are considered part of the 'takeoff' or 'landing' segments of a flight. Data is through October 31, 2001.

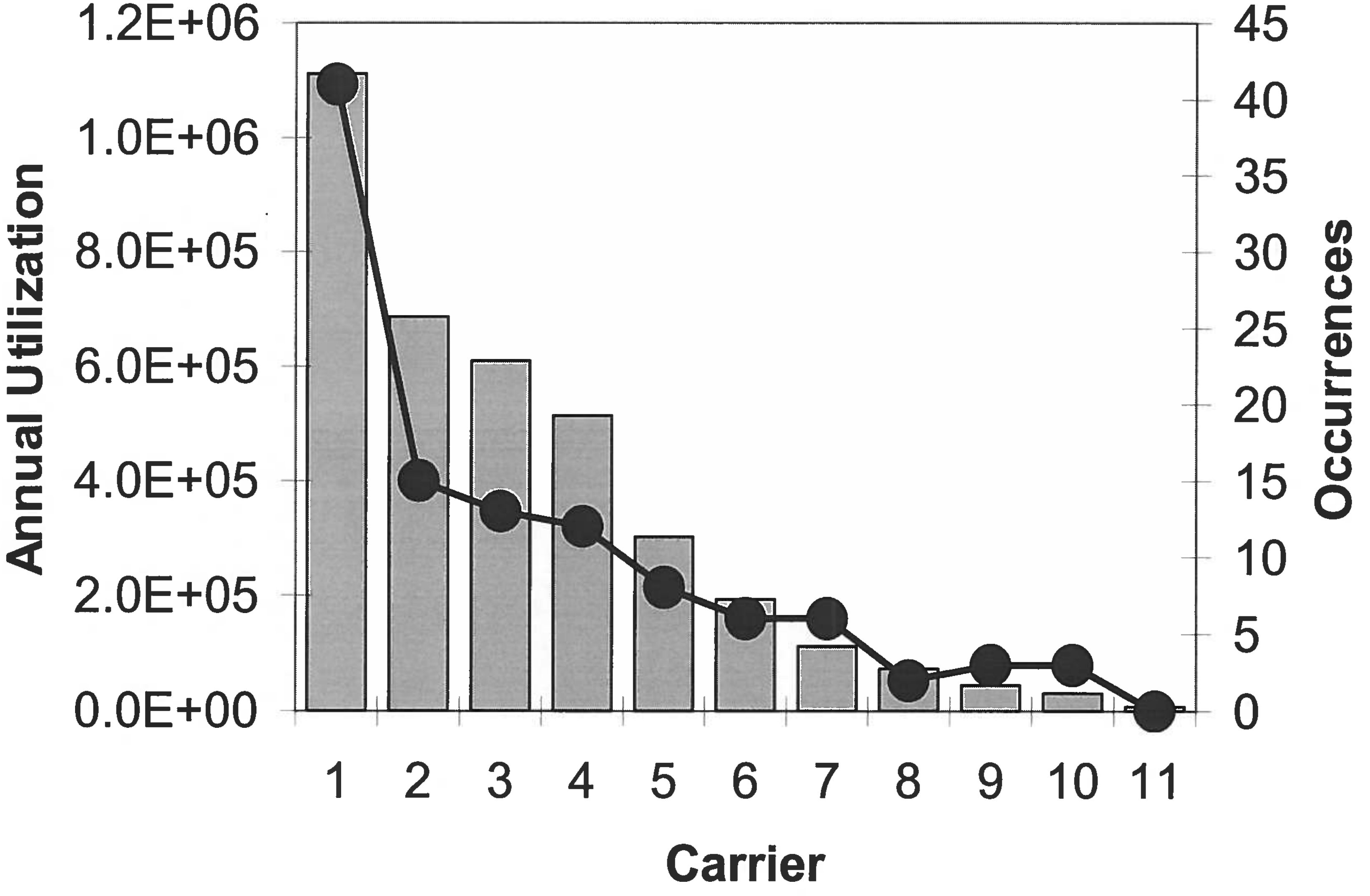
Table A.2. Aviation Event Proportions during Takeoff/Landing, Airport Operations or Enroute (Cruise)

	Year 2000			Year 2001		
Flight Segment	No.	% of Total	% of Known	No.	% of Total	% of Known
Takeoff/Landing	341	38%	45%	200	36%	47%
Enroute	291	33%	38%	170	31%	40%
Airport Operations	133	15%	17%	56	10%	13%
Unknown	121	14%		131	24%	

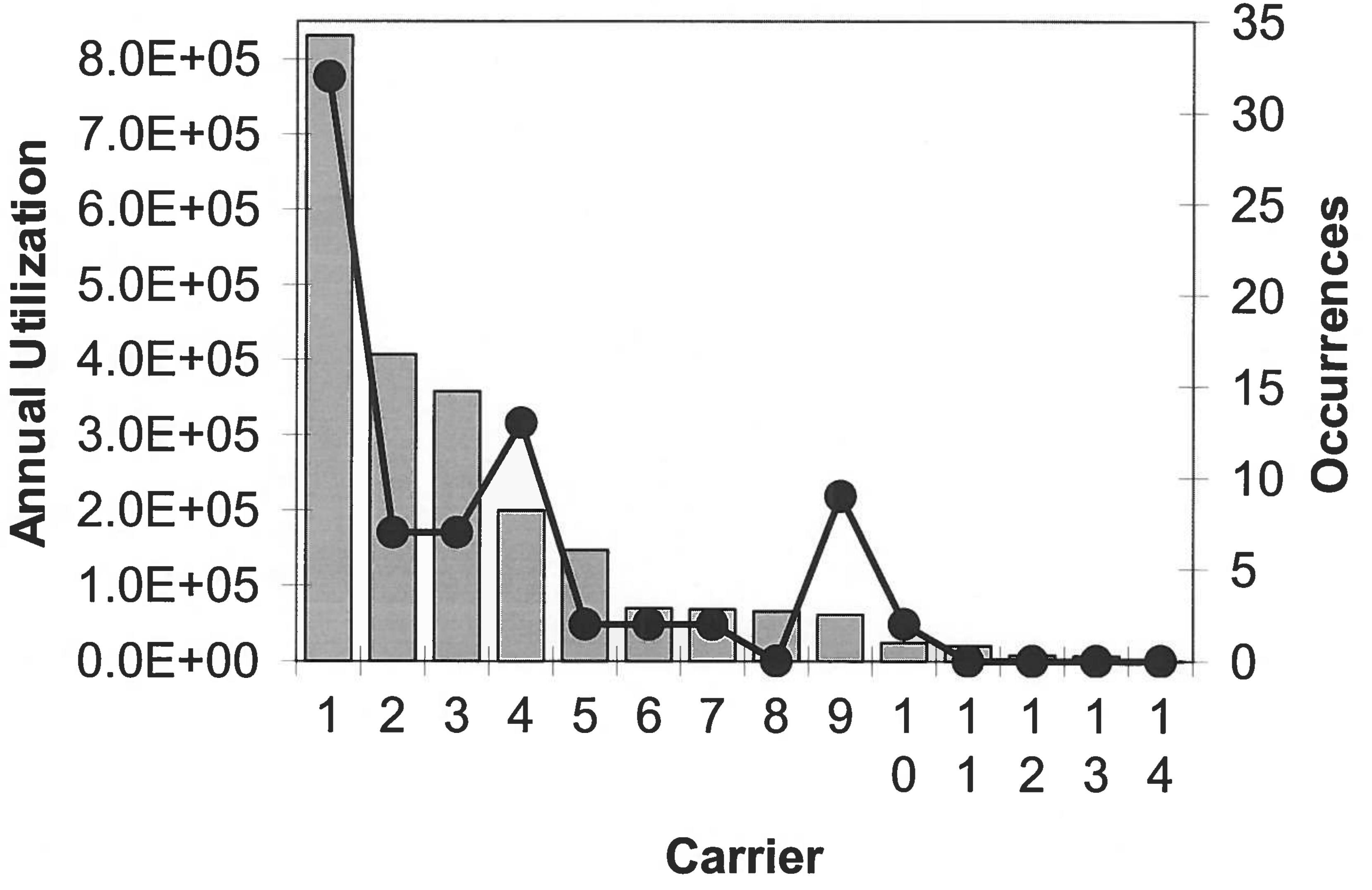
Figure 1. Histograms of Utilization versus ATA Chapter 32 System Failures (Type I aircraft in 2000). The number of year 2000 occurrences (events) due to landing gear failures for eleven air carriers is plotted against these carrier's annual flight hours in the Type I aircraft. Annual events (vertical bars) are shown against the left-hand scale and annual utilization (marked line) is shown against the right-hand scale. Even with the landing gear – active only during very early and late stages of flight – these patterns suggest the assumption that events occur proportionally to aircraft flight hours is a plausible one.

Figure 2. Histograms of Utilization versus ATA Chapter 21 System Failures (Type II aircraft in 1998). The number of year 1998 occurrences (events) due to air conditioning system failures for seven air carriers is plotted against these carrier's annual flight hours in the Type II aircraft. Annual events (vertical bars) are shown against the left-hand scale and annual utilization (marked line) is shown against the right-hand scale. Again, assuming events occur in proportion to utilization is plausible, particularly for ATA Chapter 21 and like systems, operated for the duration of every flight.

Landing Gear Events vs. Hours



Air Conditioning Events vs. Hours



Author Footnote

Lawrence Berk is an Operations Research Analyst, Aviation Safety Division, John A. Volpe National Transportation Systems Center, 55 Broadway, Cambridge, MA 02142 (Email: berkl@volpe.dot.gov). The author gratefully acknowledges the comments and suggestions of several FAA inspectors with aircraft systems expertise, particularly William Banks, John Krepp and Ross Elmore. Comments from Dr. James N. Hallock of the Volpe Center, which improved the quality of this paper, are also greatly appreciated