THE EFFECT OF SERVICE QUALITY AND EXPECTATIONS ON CUSTOMER COMPLAINTS*

SILKE J. FORBES†

Customer complaints measure consumers' dissatisfaction with the quality of a product or service. If product quality is unobservable *ex ante*, customer complaints may be driven by expectations as well as by the actually experienced quality level. I test whether the level of quality that could be expected prior to consumption affects the number of customer complaints after controlling for – *ex post* observable – actual quality, using data from the U.S. airline industry. I find that there are fewer complaints when actual quality is higher. Controlling for actual quality, a higher level of expected quality leads to more complaints.

I. INTRODUCTION

CUSTOMER COMPLAINTS ARE A MEASURE OF DISSATISFACTION with the quality of a product. In addition to providing feedback to the seller, in some contexts customer complaints can be used to signal the quality of a seller and her products to future buyers. This function of complaints or negative feedback has been analyzed extensively by authors who have studied online markets, such as eBay auctions.[1] In this paper, I study a setting in which consumers can observe information on the past quality of a firm's service, as well as the number of customer complaints that the firm received. Since past quality is directly observable, complaints are not needed to signal past quality. As a result, this setting provides an opportunity to study other effects that may influence the number of customer complaints.

I test whether, controlling for actual quality, prior expectations influence consumers' tendency to complain. There are two theoretical reasons for why expected quality might influence the number of complaints. First, behavioral models with reference-dependent utility, such as Tversky and Kahneman [1991], suggest that a consumer's utility depends not only on the actual product quality that was received but also on whether the quality was above or below a reference level, such as the product quality that the consumer expected to receive. For a given level of actual product quality a firm would receive more complaints if consumers expected to receive a higher level of quality. Second, customers may file complaints in order to penalize poor performance and encourage firms to improve the quality of their service. The discrepancy

* I would like to thank Nancy Rose and Glenn Ellison for their advice, Richard Schmalensee, participants at the 2003 Berkeley-Stanford IO Fest, the editor, Kenneth Hendricks, and two anonymous referees for helpful comments and suggestions, and the Alfred P. Sloan Foundation for financial support through the MIT Global Airline Industry Program.
† Author's affiliation: Department of Economics, University of California at San Diego, 9500 Gilman Drive, 0508, La Jolla, CA 92093-0508, U.S.A.
email: sjanuszewski@ucsd.edu

between actual quality and predicted quality may be a good proxy for the instances in which poor performance is due to mistakes made by the firm itself, as opposed to cases in which factors not under the firm's control are the cause of poor performance. In this case, we will also observe that, controlling for actual quality, the number of customer complaints is increasing in the predicted level of quality.

I study this question in the context of the U.S. airline industry with data collected by the Department of Transportation (DOT). The DOT collects and publishes data on flight delays and on the number of bags that are mishandled by each airline, and it also publishes how many complaints the DOT receives about each airline in both of those categories of service quality. Even though airline passengers cannot observe *ex ante* which level of quality they will receive, they can form an expectation of the level of quality that each airline will provide based on published information about the airline's past quality. Quality becomes observable *ex post* and passengers can decide to file complaints. I test the hypothesis that prior expectations influence passengers' tendency to complain about a given quality level by estimating the number of customer complaints per passenger that an airline receives per month as a function of its level of service quality in that month and a predicted value of consumers' prior expectations of the quality they would receive. Since I cannot observe passengers' expectations directly, I use an econometric model to predict the level of service quality for each airline-month and assume that the passengers' expectation is equal to the predicted value. I estimate several different prediction models and my results are robust across all these models.

I find, first, that the number of complaints decreases with the level of actual service quality. Second, controlling for actual service quality, consumers complain more often when they would have expected to receive higher quality. Decomposing the difference between actual and expected quality into months in which consumers received better-than-expected quality and months in which they received worse-than-expected quality, I find that worse-than-expected quality has a strong effect on the number of complaints, whereas there is no significant effect of better-than-expected quality on complaints. The results are robust across a variety of different specifications.

The size of the estimated effects of both actual and expected performance is quite large relative to the average number of complaints received. For example, the mean number of complaints about flight problems increases by 7.0 percent for a one percentage point increase in delayed flights and decreases by 5.7 percent for a one percentage point increase in expected flight delays, based on a specification that controls for carrier fixed effects and for a time trend. The mean number of complaints about baggage handling increases by 17.9 percent for each additional mishandled bag per thousand passengers and decreases by 11.3 percent for each additional expected mishandled bag per thousand passengers.

Assuming that a constant fraction of dissatisfied passengers complain each period, my findings indicate that for a given level of expectations passengers are more satisfied when they receive higher quality. In addition, for any given level of quality passengers are more likely to complain if they had higher expectations of the quality they would receive.

The results provide insights which may be important for interpreting customer feedback in settings in which past quality does not becomes directly observable, such as online marketplaces. In those cases, buyers' motivations for giving feedback about sellers may be driven by a mix of intentions to provide a signal to future buyers and of behavioral motivations. Interpretations of customer feedback in such markets may therefore need to pay attention to both of these explanations of consumer behavior.

The remainder of the paper is organized as follows: Section II presents theoretical considerations and existing evidence on customer complaints. Section III describes the data used in this study. Section IV presents the empirical model and discusses the identifying assumptions that are made in the estimation. Section V presents the results and section VI concludes.

II. Theoretical Considerations and Existing Evidence

In markets in which past product quality is unobservable to future consumers, customer complaints can mitigate adverse selection and moral hazard on the part of the seller by influencing the seller's reputation for quality [Kreps and Wilson, 1982, Milgrom and Roberts, 1982, Diamond, 1989]. When customer complaints are published and become known to future buyers, they function as a signal of the seller's quality and decrease the seller's reputation. This will lower the expected quality to future buyers and reduce their willingness-to-pay. As a result, sellers have an incentive to provide high quality to current buyers in order to avoid a negative effect on their reputation in the future.

For example, the online auction site eBay allows buyers to post positive and negative feedback about sellers. This can help future buyers to assess the quality of a seller in an otherwise anonymous online market. EBay's feedback mechanism and its effect on future sale prices has been studied empirically by a large number of authors.[2] Most of these studies find that positive customer feedback increases the price received by a seller in future transactions whereas negative feedback decreases that price, controlling for other observable factors.

In markets in which product quality is unobservable *ex ante* but becomes public knowledge *ex post*, however, past product quality can be directly observed and can be used by consumers to form expectations of future quality. In the case studied here, the uncertainty over past product quality is resolved at the same time as the information about customer complaints is published.

In this setting, one possible explanation for why consumers file complaints is provided by a behavioral model. Tversky and Kahneman [1991] develop a model in which consumers have utility functions that are dependent on a reference level. In their work, the reference level is usually thought of as the status quo, such as the current consumption or the initial endowment, but the authors suggest that the reference level can also be influenced by consumers' expectations. Assume that at the time of purchase – before the actual consumption experience – consumers form an unbiased expectation of the level of the quality that they will receive. If this expectation becomes the reference level against which the actual consumption experience is compared, then consumers will be satisfied if they receive the expected quality or better, and will be dissatisfied – and tend to complain – otherwise. Only a fraction of consumers who are dissatisfied complain. This is likely due to the costs associated with filing a complaint, which vary across consumers. If one assumes that these costs are independently and identically distributed across consumers and are orthogonal to the quality of service received, then the fraction of consumers that complain will be constant across periods.

An alternative explanation for why customers file complaints in this setting is that poor performance may either be due to mistakes by the firm or may be caused by factors outside the firm's control, and customers intend to penalize the first but not the latter. In the airline industry, flight delays and mishandled baggage can be caused, on the one hand, by bad weather or air traffic control problems and, on the other hand, by poor organization or maintenance failures on the part of the airlines. Under this explanation, if airline passengers are able to identify the cause of flight delays or mishandled baggage, they will complain about these only when they are due to the airlines' mistakes. As researchers, we are not able to observe the cause of poor service quality but the predicted level of quality may provide a good proxy for the extent to which flight delays and mishandled baggage in any given period are due to factors outside the airlines' control. Alternatively, airline passengers may be unable to identify the causes of poor quality but may themselves use expected quality as a proxy for the level of quality that they should receive absent any mistakes on the part of the airlines.

Under either of these models, the number of complaints will be decreasing in the level of actual quality received and increasing in the level of quality that was expected prior to the consumption experience.[3] Another implication of the model is that the number of complaints depends on the deviation of the actual quality from the expected quality. In particular, worse-than-expected performance will increase the number of complaints. I will test both of these implications empirically.

Similar models have been examined by some of the marketing research on customer satisfaction [e.g. Oliver, 1980, and the references listed in Oliver and DeSarbo, 1988]. This work assumes that a consumer's satisfaction with a product or a service depends on the actually experienced quality as well as on the expected quality. In empirical tests, a number of studies have found that the deviation of actual quality from expected quality affects customer satisfaction [Oliver, 1980, and the references cited therein, p.461, Bearden and Teal, 1983, Oliver and DeSarbo, 1988, Anderson and Sullivan, 1993, de Ruyter *et al.*, 1993, Johnson *et al.*, 1995]. However, these studies rely on survey data to measure satisfaction, actual quality, and expected quality. Often, the survey data are collected after the consumption experience which may lead to reporting biases in all these variables. The rational expectations approach employed here is new to the literature on customer satisfaction and complaints. Nevertheless, my findings are consistent with results from previous studies. It is worth noting, however, that in the case studied here customers can only express dissatisfaction but not positive feedback. This may also be important when comparing my findings to studies of eBay's feedback mechanism because on eBay most submitted feedback is positive, and negative feedback is quite rare.[4]

Another possible explanation for why customers complain in the setting studied here could be that they merely file complaints because they expect to receive a direct benefit, such as monetary compensation. Empirically, this would imply that the number of complaints is decreasing in the actual quality of service, since any consumer who received poor quality would want to complain. Prior expectations would not affect the number of complaints. However, the DOT explicitly states that there are no direct benefits to passengers who file complaints with it. For example, the DOT does not assist in complaint resolution and does not give the airlines the names of passengers who have filed complaints. As I will show in section 5, this explanation for customer complaints can be rejected since I find that prior expectations have a statistically significant effect on the number of complaints.

III. DATA

This study uses data for the airline industry which are collected by the U.S. Department of Transportation (DOT). Airline passengers can register air travel service problems with the DOT, independent of whether they have already filed a complaint with the airline itself or not. The DOT publishes the number of complaints it receives in its monthly Air Travel Consumer Report.

The report aggregates complaints by airline and by the month in which the complaint is filed. The majority of complaints concerns incidents that occurred in the same month or one month prior to the filing of the complaint. There is no direct benefit to the passenger from filing a complaint with the DOT. In particular, the DOT does not mediate individual customer complaints.

The complaints are registered in eleven different categories. Among these categories are flight problems, which encompass delays, cancellations and other schedule disruptions, and baggage, for claims concerning lost, damaged or delayed baggage. For these two categories, monthly data on each airline's performance are also available in the Air Travel Consumer Report.[5] The reported statistics are the number of mishandled – i.e. lost, damaged, or delayed – bags per thousand passengers and the percentage of flights delayed more than fifteen minutes. As with the complaint data, the performance data are aggregated by month and carrier for all routes. In addition, the performance data are reported only for major airlines. These are all U.S. airlines with at least one percent of total domestic scheduled-service revenues.

Since the data are only reported as aggregated to airline-months, I will test the effect of average observed performance and average predicted performance on the number of complaints in my empirical estimation. I am not able to test the model at the level of an individual consumer.

The DOT has collected and published the information on passenger complaints, on-time performance and mishandled bags since 1988. This study includes data from January 1988 to September 2000 for seven major carriers: American Airlines, Continental Airlines, Delta Airlines, Northwest Airlines, TWA, United Airlines, and US Airways.[6] This results in 1071 carrier-month observations. After eliminating observations for which either the performance measures or the number of complaints are missing, I have 993 observations for flight delays and 1001 observations for mishandled bags left in my sample.

Since 1995, more detailed, flight-level data on arrival times are available. In that year, the DOT also started to include delays that are due to maintenance problems in its reported statistics. I therefore perform robustness checks using (i) only flight delays after 1995, and (ii) an alternative measure of flight delays, the percentage of flights delayed over 45 minutes.[7] Flight cancellations have only been reported since the second half of 1998. I do not include them in my estimation.

Both the aggregated and the detailed on-time arrival data report *flight*-level delays which may deviate from the average delays experienced by passengers. Bratu and Barnhart [2004] study proprietary data provided by one major airline and find that flights which transport a greater number of passengers are on average more delayed. In addition, the flight-level average does not account for passengers who miss connecting flights. Bratu and Barnhart demonstrate in simulations that the statistics published by the DOT tend to underestimate the true delays experienced by passengers. However, they also show that the published on-time statistics are positively correlated with the true delay so that the DOT statistics can be used as proxy for the true experienced delay.

Table I presents summary statistics on complaints and on service quality. The mean number of complaints about flight problems per million passengers is 6.9 per month, with a standard deviation of 10.5. The minimum is zero and the maximum is 147.8. Complaints are censored at zero.[8] I observe zero complaints for 7 out of 993 observations in the sample. The mean percentage of flights delayed over fifteen minutes is 21.8, with a minimum of 6.9 and a maximum of 63.9 during Northwest's pilot strike in September 1998. The standard deviation of this variable is 6.9.

For the category of mishandled baggage, the number of complaints per million passengers is on average 3.0 per month with a standard deviation of 4.1. I observe zero complaints for 23 of 1001 observations. The maximum of this variable is 54.1. On average, 5.9 bags per thousand passengers are mishandled, with a minimum of 2.9 and a maximum of 19.0. The standard deviation is 1.8.

Place Table I approximately here

Table II shows means and standard deviations by airline. There is substantial variation in complaints across airlines. Delta Airlines receives the fewest complaints per passenger in both categories, whereas TWA receives the most complaints per passenger, also in both categories. The standard deviations of complaints, reported in parentheses, are quite large indicating a substantial amount of variation in complaints within airlines over time. Flight delays and the number of mishandled bags also vary across airlines, but much less so than the complaints. Northwest has the fewest flight delays with an average of 19.6 percent of its flights delayed, while United Airlines has the most delayed flights with 24.3 percent on average. For mishandled bags, we observe the best performance by Continental with 5.3 mishandled bags per thousand passengers, and TWA has the most mishandled bags per thousand passengers with an average of 6.6.

Place Table II approximately here

IV. Empirical Model and Identification

I model the number of customer complaints received by firm i in period t as a function of the actual (observed) performance of the firm in that period, s_{it}, and its expected performance in that period, $E(s_{it})$, allowing for a constant effect β_0. In some of the specifications, I also include a time trend and carrier fixed effects.

$$(1) \qquad Complaints_{it} = \beta_0 + \beta_1 s_{it} + \beta_2 E(s_{it}) + \varepsilon_{it}$$

This specification allows me to test the effect of actual quality and of expected quality separately. In an alternative specification, I test the hypothesis that the deviation of actual quality from expected quality affects the number of customer complaints by imposing the restriction that $\beta_1 = -\beta_2$. Finally, I estimate separate coefficients for positive and negative deviations of actual quality from expected quality in order to test whether worse-than-expected quality and better-than-expected quality have effects of different magnitude. The theoretical model predicts that better-than-expected quality will have no effect on complaints if consumers are dissatisfied only if the experienced quality is worse than expected.

The estimation proceeds in two steps. First, I estimate the relationship between an airline's performance and observables and use that model to predict an expected value of quality, $E(s_{it})$, for each firm in each time period. In the second step, I estimate the model for customer complaints in equation (1) using the predicted values from the first step. The remainder of this section details the estimation of both steps and discusses identification.

In order to generate expectations for on-time arrival and baggage-handling, I estimate an econometric model for both variables and assume that the predictions from this model are equal to the passengers' expectations. In order to check the robustness of the prediction, I estimate several alternative models for the passengers' expectations. The baseline model estimates the percentage of delayed flights and the number of mishandled bags per thousand passengers as functions of month and airline fixed effects. The model assumes that passengers form their expectations for flight delays and mishandled bags only based on information about the month in which they are flying and the carrier that they are flying with. An even simpler model predicts flight delays and the number of mishandled bags only with month fixed effects, not including airline fixed effects. This is the second model that I estimate. This model does not require passengers to form airline-specific expectations. Instead, their expectations are only based on seasonal effects.

More sophisticated models of passengers' expectations can include information about strikes, bankruptcies, and the size of the airline's operations such as the number of departures. I report results of a model which estimates expected quality as a function of month and airline fixed effects as well as dummy variables for the months in which an airline experiences a strike or operates in bankruptcy protection.[1] Periods of bankruptcy and strikes are rare but can severely disrupt operations. These period can often be predicted in advance. Note, however, that in this specification the predicted value $E(s_{it})$ contains the passengers' prediction error of whether a strike or a bankruptcy will occur. Unbiased estimation of the parameters of the model requires that this prediction error is orthogonal to the other covariates.

In the second step of the estimation, I regress the number of customer complaints per million passengers in the categories 'flight problems' and 'baggage' on the actual and the predicted values of the percentage of delayed flights and the number of lost bags per thousand passengers, respectively. I use Tobit regressions to account for the censoring of complaints at zero. The standard errors are bootstrapped.

In the baseline regressions, I include only the actual and the predicted value of the carrier's performance in the time period. I also report alternative specifications which include a time trend. This time trend controls for possible variations over time in passengers' propensity to complain. Such variations may occur, for example, due to unobserved changes in the costs of filing a complaint. In addition, there may be awareness effects, for example in periods in which news media provide extensive coverage of service quality problems in the airline industry. Finally, I estimate a specification that controls for carrier fixed effects as well as for the time trend. Carrier fixed effects may have an effect on the number of complaints received by a carrier even after controlling for actual and expected quality because carriers may handle incidences of flight delays and mishandled baggage differently. For a given level of actual and of expected quality, carriers with better customer service helping out passengers who are delayed or whose

[1] A previous version of the paper included estimations using information on the airline's number of departures and number of passengers in each month. The results are available from the author upon request. Note that when month and carrier fixed effects are included in the regression, any additional effect of the size of the airline's operations will come from changes in those variables over time. The results reported here are robust to including those additional variables, as well as to estimating carrier-specific slopes for these variables and for the effects of strikes and bankruptcies.

bags are delayed or lost will get fewer complaints from passengers than other carriers with worse customer service.

The identifying assumption for the estimation of the effect of expected quality on passenger complaints is that the month fixed effects and strike and bankruptcy periods affect passenger complaints only through expected quality and have no direct effect on the number of complaints. If this assumption were violated, the predicted value of quality could pick up the direct effect of these variables, which would lead to a biased coefficient on expected quality. This identifying assumption is justified by the process of filing complaints with the DOT. Passengers have the same access to the DOT for filing complaints regardless of the month that they flew in. This access also does not change during periods in which some airlines experience a strike or are in bankruptcy protection. The cost of filing a complaint should therefore not be affected by any of these variables. The specifications which include a time trend allow for continuous changes in the ease of filing a complaint or for awareness effects as the ones described above. We should not expect these factors to be captured well by month dummies. Likewise, the benefit which passengers derive from filing a complaint should not vary by calendar month.

In the baseline specification for passenger complaints which does not include carrier effects, the identifying assumption for the carrier effects is that any systematic differences across carriers in the number of complaints received are orthogonal to the effects of actual and expected quality. In an alternative specification, I allow carrier fixed effects to affect the number of complaints directly, as well as through the passengers' expected quality, in order to allow for differences in customer service across airlines which may affect the passengers' propensity to complain. The results in section 5 show that these carrier effects are statistically significant, and the coefficients on actual and expected quality are robust to the inclusion of these fixed effects.

V. RESULTS

I start with the presentation of the results for the first step of the estimation which is used to predict expected quality. I then present the second step of the estimation in which I estimate the number of complaints as a function of actual and expected quality.

V(i). *Predicting Expected Quality*

I estimate quality in the two categories of interest, flight delays and mishandled bags, as a function of observables. This model is then used to predict expected quality for each airline-month. I present several different specifications of the model.

I begin by showing a variance decomposition for the baseline model with month and airline fixed effects to investigate the relative importance of these exogenous sources of variation in quality. Table III shows in Panel A that for flight delays the month fixed effects explain 22.6 percent of the total sum of squares while the carrier fixed effects explain only 3.4 percent of the total. Panel B shows the variance decomposition for mishandled bags. Here, month fixed effects explain 30.5 percent of the total sum of squares and carrier fixed effects explain 6.6 percent of the total. If we were to include fixed effects for each time period instead of the month fixed effects, these would explain 57 percent of the total variation in flight delays and 67 percent of the total variation in mishandled bags. This shows that most of the variation in our measures of service quality comes from the time series rather than across carriers.

Place Table III approximately here

Table IV, Panel A presents the regression results for flight delays. The first column shows a regression of the percentage of flights delayed more than 15 minutes on month fixed effects only. The omitted month dummy is for January. I find that the most flight delays occur in the winter months, especially in December and January. The fewest delays are found during the fall months. The difference in the percentage of delayed flights between January, the month with the most delays, and September, the month with the fewest delays, is 10.7 percentage points. This is almost half of the mean value of delayed flights. All month dummies have a negative coefficient and all except the one for December are statistically significant at the 1 percent level. The R-squared for this model is 0.23.

Place Table IV approximately here

Column 2 adds carrier fixed effects to the specification. The omitted carrier is American Airlines. This is the baseline model which I use to predict passengers' expectations of flight delays. I find that adding the carrier fixed effects does not lead to statistically significant differences in the estimated coefficients for the month effects. Many of the carrier dummies are not individually significant. Only Northwest Airlines has significantly fewer delays than American Airlines, and United Airlines has significantly more delays than American Airlines. However, an F-test of the joint significance of the carrier effects rejects the null hypothesis that all of the carrier coefficients are equal to zero. The R-squared of the model increases from 0.23 in column 1 to 0.26 here, reflecting the finding from the ANOVA table that most of the variation in flight delays is explained by month effects rather than by carrier effects.

The last column of Panel A adds dummy variables for strike and bankruptcy periods as further explanatory variables. These are equal to one for periods in which the carrier experiences a labor strike or is in bankruptcy protection, respectively. The carriers that experience strikes at some point during the sample period are American Airlines, Northwest Airlines, and US Airways. Carriers that were in bankruptcy protection for some of the sample period are American Airlines, Continental Airlines, and TWA. I find a positive but statistically insignificant point estimate on the effect of strikes. For carriers in bankruptcy protection, I find a statistically significant reduction in flight delays of 3 percentage points. When carriers enter bankruptcy protection, they typically reduce the number of flights that they operate, decreasing the degree of congestion in their network. This is likely to explain the reduction in flight delays that occurs during bankruptcy protection. The point estimates for all of the carrier effects increase in magnitude when I add the controls for strikes and bankruptcies and the carrier effects remain jointly significant. The positive coefficients on Continental and Delta are now statistically significant at the 10 percent level. The R-squared of the estimation increases to 0.28.

Even though the fit of the estimation is slightly higher for the specification in column 3 than for column 2, I use the results from column 2 as the baseline model for predicting passengers' expectations of quality. I do this because that specification requires less information on the part of the passengers. However, I do report robustness checks of my later results using the model from column 3 as well as the simpler model from column 1 to predict expectations. My findings on the effect of expectations on customer complaints are robust to using either of these models. Further robustness checks include models that estimate carrier-specific slopes for all month effects, with and without including the strike and bankruptcy controls. I also estimate the model including a control for the total number of departures of the carrier during the time period. Those results are not reported but are available upon request. All these models produce predicted values for passenger expectations which lead to qualitatively the same conclusions on the effect of expectations on customer complaints as the models that I present here.

Panel B of Table IV presents the same regressions for mishandled bags. The dependent variable is the number of mishandled bags per thousand passengers. In the first column, which includes only month dummies as explanatory variables, I find a similar pattern as for flight delays. The number of mishandled bags is highest during the winter months, especially in December and January, and lowest during the fall. The difference between December, the month with the most mishandled bags, and September, the month with the fewest mishandled bags, is estimated to be 3.3 bags per thousand passengers which is more than half of the mean value of mishandled bags in the sample. 10 out of 11 month dummies have a negative coefficient and all except the one for December are statistically significant at the 1 percent level. The R-squared for this model is 0.30.

Column 2 adds carrier fixed effects. I find that Northwest, TWA, and United Airlines have significantly more mishandled bags than the omitted carrier, American Airlines. None of the other carriers show individually significant differences to American Airlines. The test of joint significance of the carrier effects strongly rejects the null hypothesis that all of the coefficients are equal to zero. Again, none of the coefficients on the months are significantly affected by the inclusion of the carrier effects. The R-squared increases from 0.30 in column 1 to 0.37 here. Finally, I add the controls for strike and bankruptcy periods. I find that neither of these effects can be distinguished from zero. The coefficients on the carrier and month dummies are also not significantly affected. Again, I use the results from column 2 as my baseline for predicting passenger expectations in the next section, and I present robustness checks using the models from columns 1 and 3.

V(ii). *Results on Customer Complaints*

I estimate the effect of the actual quality of service and of the predicted value of quality from the previous section on the number of complaints per million passengers. The estimation uses Tobit regressions to account for the censoring of the dependent variable at zero. All reported standard errors are bootstrapped because the expected quality is a predicted value from another estimation. Table V, Panel A reports the first set of results for the category of flight problems. I start in column 1 by showing the effect of the actual percentage of flight delays on complaints without any other controls. I find a positive relationship with a highly significant point estimate of 0.451. The interpretation is that complaints become more common when actual quality is worse, i.e. when flight delays are more frequent.

Place Table V approximately here

In column 2, I present the test of the main hypothesis of the paper which is that expected quality as well as actual quality affect customer complaints. I regress the number of customer complaints on the actual percentage of flight delays and the predicted percentage of flight delays, using the model in column 2 of Table IV to generate the prediction. I find that there are more complaints when actual flight delays increase, and that there are *fewer* complaints when the *expected* percentage of flight delays increases. My estimates imply that passengers' expectations have a significant effect on their propensity to complain about the quality of service which they receive. In particular, passengers are less likely to complain about any given level of quality when they expected to receive lower quality. The point estimates in the model are 0.581 for actual flight delays and -0.501 for expected flight delays. This implies that a one percentage point increase in actual flight delays increases complaints by 0.581 per one million passengers, whereas a one percentage point increase in expected delays decreases complaints by 0.501 per one million passengers. An increase in flight delays by one standard deviation will increase the number of complaints by 4.0 per 1 million passengers or 0.38 standard deviations. Both effects are statistically significant at the 1 percent level.

Column 3 checks these results for robustness by controlling for a quadratic time trend. I find that the coefficients on actual flight delays as well as on expected flight delays are robust to including this time trend. I find a negative coefficient on the linear term of the time trend and a positive coefficient on the quadratic term. The implied pattern is that complaints decreased for the first part of the sample period and increased thereafter. In column 4, I include carrier fixed effects in addition to the time trend. The omitted carrier is American Airlines. The coefficients on actual delays, expected delays, and the time trend are barely affected by the inclusion of carrier effects. All remain statistically highly significant. The carrier effects are jointly significant, and are individually significant for four of the carriers. Controlling for actual and expected quality, Continental, Northwest, and TWA receive more complaints than American Airlines, and Delta receives fewer complaints than American. This is likely due to differences in the customer services response of these airlines. As the results in Panel B will show, I find a similar pattern across carriers for complaints about mishandled baggage.

Finally, I test whether the effects of actual and expected quality are robust to using a different empirical model to predict passengers' expectations. I first use the more naive model from column 1 in Table IV in which passengers form expectations about flight delays purely based on the month of the year in which they are flying. This model does not require passengers to be able to distinguish between airlines. The results reported in column 5 are for a specification without time and carrier effects, but are robust to including those additional controls. I still find that complaints increase significantly with actual flight delays and decrease significantly with expected flight delays. The point estimates are smaller than the estimates in the comparable specification in column 2, but are not statistically different from those. The same is true for the other robustness check which is presented in column 6. Here, I use the more sophisticated model for passenger expectations which predicts flight delays based on the month, the carrier, and the controls for periods with strikes and bankruptcies. I again find slightly smaller point estimates than in column 2, but both effects are not statistically different from the estimates in column 2.

Panel B of Table V presents the same estimations for the other observed category of service quality, mishandled baggage. Again, I begin in column 1 by presenting the effect of actual quality on customer complaints, without including any other controls. Here, I also find that complaints become more frequent as quality decreases. The dependent variable in this regression is the number of customer complaints about mishandled bags per million passengers, and the variable that measures actual quality is the number of mishandled bags per thousand passengers. My estimate implies that an additional mishandled bag leads to an additional 0.766 complaints per thousand passengers. This relationship is statistically highly significant.

In column 2, I test the main hypothesis – that actual quality as well as expected quality affect complaints – for this second category of mishandled bags. I include the actual number of mishandled bags as well as the predicted number of mishandled bags as explanatory variables. The prediction is again based on the specification with month and carrier effects (presented in column 2 of Table V, Panel B). I find, as before, that the number of complaints increases when service quality decreases, i.e. when more bags are delayed or lost, but that controlling for actual quality, complaints are less frequent when *expected* quality decreases. The estimated coefficients are 0.952 on actual mishandled bags and -0.505 for the expected number of mishandled bags. An increase in mishandled bags by one standard deviation leads to 1.7 additional complaints per million passengers.

When I add a quadratic time trend to the model in column 3, the signs on actual and expected mishandled bags remain the same but the both effects are estimated to be smaller than before and the coefficient on expected quality becomes statistically insignificant. This result may be due to a poor fit of the prediction model. When I exclude the carrier with the most volatile performance, TWA, from the sample and re-estimate the equation I find a negative and statistically significant effect of expected quality in specifications without and with the time trend. When I include carrier fixed effects in addition to the time trend in column 4, there is little effect on the coefficient on actual quality compared to column 3. The coefficient on expected quality increases in magnitude to -0.336 and is now again statistically significant. For the time trend itself, I find a negative coefficient on the linear term and a positive coefficient on the quadratic term. The estimates imply that complaints decreased throughout the sample period. The carrier fixed effects in column 4 are jointly significant and are individually significant for four of the carriers. The pattern of the individual carrier effects is similar to the one found for complaints about flight delays which suggests that there are systematic differences across airlines in the handling of disruptions which lead to the differences in passengers' propensity to complain about these airlines.

In columns 5 and 6, I again check the results for robustness using the alternative models for predicting passengers' expectation of quality. First, I use the measure of expected quality which is predicted only based on month effects (from column 1 in Table IV, Panel B). I continue to find a positive effect of actual mishandled bags on complaints and a negative effect of the expected number of mishandled bags, with coefficients that are slightly larger in magnitude than in the comparable specification in column 2. The second alternative model of expectations, which forms predictions based on month and carrier effects as well as controls for periods in which the carrier experiences a strike or is in bankruptcy protection, produces results which are very close to the ones in column 2.

Next, I investigate whether complaints are affected more strongly in periods when actual quality is worse than expected than when actual quality is better than expected. I first show the effect of the prediction error – the difference between actual and expected quality – on complaints and then decompose the effect into worse-than-expected performance and better-than-expected performance. The measure of expectations that is used here is the one based on month and carrier effects. The findings are robust to using the other measures of expectations.

Table VI presents the results of this second set of regressions. Panel A starts with the findings in the category of flight problems. In column 1, I regress the number of complaints per million passengers on the difference between actual and expected flight delays. I find a significantly positive effect with a point estimate of 0.581. This effect is robust when I include the time trend, as reported in column 2, and carrier fixed effects as well as a time trend, as reported in column 3.

Place Table VI approximately here

When I decompose the difference between actual delays and expected delays into a positive and a negative deviation, the positive deviation, which reflects worse-than-expected quality, is estimated to have a larger effect of 0.901 and it is highly significant. The negative deviation, in contrast, has a much smaller coefficient of 0.032 and is not statistically different from zero. This pattern persists when I include the time trend and carrier effects in columns 5 and 6. The interpretation of these findings is that better-than-expected performance does not reduce the number of passenger complaints, whereas worse-than-expected performance increases the number of complaints by a large and significant amount. According to these estimates, a carrier whose flights are delayed more often than expected by 6.9 percentage points, or one standard deviation of flight delays, will receive 6.2 more complaints per million passengers which is not much less than the average number of complaints per million passengers.

I find very similar conclusions for mishandled baggage. These results are shown in Panel B of Table VI. I find a coefficient of 0.952 for the deviation of actual quality from expected quality in column 1. This coefficient decreases to 0.531 when I add the time trend in column 2, and to 0.529 when I add carrier dummies as well as the time trend.

When I decompose the deviation into worse-than-expected and better-than-expected performance, I find again that worse-than-expected quality has a large and statistically significant effect on the number of complaints while the effect of better-than-expected performance cannot be statistically distinguished from zero. This is true without and with a time trend and carrier effects in the model, but the magnitude of the coefficients is smaller when the time trend and the carrier effects are included. Column 6, which has the largest set of controls, shows an effect of worse-than-expected performance of 0.742. This implies that each mishandled bag per thousand passengers above the expected level leads to 0.742 additional complaints per million passengers, which is equal one quarter of the average number of complaints per month in this category. As in the case of flight delays, this effect is quite large.

Overall, I find strong support for the hypothesis that customer complaints increase as actual quality decreases, and that complaints are more likely when customers expect to receive higher quality. Worse-than-expected quality leads to a significant increase in complaints, while better-than-expected quality does not affect the number of customer complaints.

The remainder of this section describes some additional robustness checks which were performed for flight delays. The results are not reported here but are available upon request. The DOT changed its reporting of flight delays in 1995. Flights which were delayed due to mechanical problems were not included in the DOT's statistics before 1995 but were included after this date. I re-estimate the regressions for flight delays including only data from 1996 to 2000. My findings are robust to these restrictions.

In 1995, the DOT also started to collect more detailed data on flight delays which allow to construct other aggregate measures as an alternative to the DOT's summary statistic, the percentage of flights delayed over 15 minutes. I compute the percentage of flights that are over 45 minutes delayed for each carrier in each month and test the effect of actual and predicted flight delays over 45 minutes on the number of complaints. The results confirm the earlier findings that the number of complaints is increasing in the percentage of actual flight delays and decreasing in the percentage of expected flights delays.

As a final robustness check, I regressed the number of complaints in the current month on the actual and the expected service quality of the previous month. Complaints are reported by the DOT for the month in which they are filed. This includes incidents from the current month as well as from past months. The majority of these incidents are from the current and from the previous month. The results for specifications using the actual and the expected service quality from the previous month rather than from the current month are quite similar to the results presented here. None of the estimated effects change sign. The coefficient estimates are generally smaller in magnitude, indicating that the performance in the previous month has a smaller effect on customer complaints than performance in the current month.

VI. CONCLUSION

This paper estimates the effect of actual quality and of expected quality on customer complaints in a setting in which quality is unobservable *ex ante* but becomes observable and public knowledge *ex post*. I use data from the U.S. airline industry reporting the number of complaints and measures of service quality aggregated to the airline-month level for the years 1988-2000. Since I do not observe expected quality directly, I estimate an econometric model of service quality as a function of month and carrier fixed effects and assume that the passengers' expectations are equal to the predicted values from this model. I test for robustness using a variety of alternative prediction models.

I find that the number of complaints per passenger increases when actual quality decreases and that, controlling for actual quality, higher expected quality leads to more customer complaints. When I decompose the effect into better-than-expected and worse-than-expected quality, I find that worse-than-expected quality has a large effect on complaints whereas there is no statistically significant effect of better-than-expected quality on the number of complaints. These findings are interesting when compared to other industries in which the actual quality level experienced by current consumers remains unknown to future buyers, such as online auctions. In those settings, there is evidence that customer feedback serves to signal the quality of a seller or a product to future buyers. Here, past quality can be observed directly by future buyers and can be used to form expectations of future quality. In contrast to studies of eBay's feedback mechanism, however, the case studied here only allows consumers to express complaints, not positive feedback. Even though past quality becomes observable, consumers still file complaints in this industry. This may be due to behavioral explanations under which consumers complain if the quality they received was less than what they expected. Or, consumers may wish to only punish mistakes due to the airline's own behavior, and expected quality may serve as a proxy for poor quality caused by external factors not under the airline's control.

REFERENCES

Anderson, E.W. and M.W. Sullivan, 1993, 'The Antecedents and Consequences of Customer Satisfaction for Firms', *Marketing Science* 12, pp. 125-143

Bajari, P. and A. Hortacsu, 2004, ‘Economic Insights from Internet Auctions’, *Journal of Economic Literature* 42, pp. 457-486

Bearden, W.O. and J.E. Teel, 1983, ‘Selected Determinants of Consumer Satisfaction and Complaint Reports’, *Journal of Marketing Research* 20, pp. 21-28

Bratu, S. and C. Barnhart, 2004, ‘An Analysis of Passenger Delays Using Flight Operations and Passenger Booking Data’, *Sloan Industry Studies Working Papers* WP-2004-20

Darley, J.M. and R.H. Fazio, 1980, ‘Expectancy Confirmation Processes Arising in the Social Interaction Sequence’, *American Psychologist* 35, pp. 867-881

De Ruyter, K., J. Bloemer and P. Peeters, 1997, ‘Merging service quality and service satisfaction. An empirical test of an integrative model’, *Journal of Economic Psychology* 18, pp. 387-406

Diamond, D.W., 1989, ‘Reputation Acquisition in Debt Markets’, *Journal of Political Economy* 97, pp. 828-862

Johnson, M.D., E.W. Anderson and C. Fornell, 1995, ‘Rational and Adaptive Performance Expectations in a Customer Satisfaction Framework’, *Journal of Consumer Research* 21, pp. 695 - 707

Kreps, D.M. and R. Wilson, 1982, ‘Reputation and Imperfect Information’, *Journal of Economic Theory* 27, pp. 253-279

Milgrom, P. and J. Roberts, 1982, ‘Predation, Reputation, and Entry Deterrence’, *Journal of Economic Theory* 27, pp. 280-312

Oliver, R.L., 1980, ‘A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions’, *Journal of Marketing Research* 42, pp. 460-467

Oliver, R.L. and W.S. DeSarbo, 1988, ‘Response Determinants in Satisfaction Judgments’, *Journal of Consumer Research* 14, pp. 495-507

Resnick, P., R. Zeckhauser, J. Swanson and K. Lockwood, 2006, ‘The Value of Reputation on EBay: A Controlled Experiment’, *Experimental Economics* 9, pp. 79-101

Tversky, A. and D. Kahneman, 1991, ‘Loss Aversion in Riskless Choice: A Reference-Dependent Model’, *Quarterly Journal of Economics* 106, pp. 1039-1061

U.S. Department of Transportation, *Air Travel Consumer Report*, various issues

Zebrowitz, L.A., 1997, *Reading Faces: Window to the Soul?* (Westview Press, Boulder, CO).

TABLE I: DESCRIPTIVE STATISTICS

Variable	Observations	Mean	Std. Dev.	Min	Max
Complaints about flight problems per million passengers	993	6.9	10.5	0	147.8
Delays over 15 minutes, in percent	993	21.8	6.9	6.9	63.9
Complaints about mishandled bags per million passengers	1001	3.0	4.1	0	54.1
Mishandled bags per 1000 passengers	1001	5.9	1.8	2.9	19.0
Strike	1001	0.01	0.10	0	1
Bankruptcy	1001	0.05	0.22	0	1

Source: Department of Transportation, *Air Travel Consumer Report*, years 1988-2000.

TABLE II: DESCRIPTIVE STATISTICS BY CARRIER

Carrier	American Airlines	Continental Airlines	Delta Airlines	Northwest Airlines	TWA	United Airlines	US Airways
Complaints about flight problems per million passengers	6.0	9.1	2.6	8.5	10.5	6.7	5.2
	(6.1)	(16.8)	(2.4)	(13.2)	(9.8)	(10.1)	(6.3)
Delays over 15 minutes, in percent	21.3	21.7	22.2	19.6	21.9	24.3	21.7
	(6.2)	(5.2)	(5.7)	(7.8)	(7.5)	(7.8)	(6.5)
Complaints about mishandled bags per million passengers	2.7	4.0	1.2	2.5	6.3	2.5	1.6
	(2.1)	(6.3)	(1.0)	(3.4)	(6.0)	(2.3)	(1.6)
Mishandled bags per 1000 passengers	5.5	5.3	5.5	6.0	6.6	6.5	5.7
	(1.5)	(1.6)	(1.6)	(1.7)	(2.4)	(1.7)	(1.9)

Mean values. Standard deviations in parentheses.
Source: Department of Transportation, *Air Travel Consumer Report*, years 1988-2000.

TABLE III: VARIANCE DECOMPOSITION (ANOVA) FOR MODEL OF PASSENGER EXPECTATIONS

PANEL A: DEPENDENT VARIABLE IS PERCENTAGE OF DELAYED FLIGHTS

Source	Partial Sum of Squares	Degrees of freedom	Ratio of Partial to Total Sum of Squares
Month	10545.4	11	0.226
Carrier	1607.7	5	0.034
Total	46608.0	790	1.000
R-squared	0.261		

PANEL B: DEPENDENT VARIABLE IS MISHANDLED BAGS PER 1000 PASSENGERS

Source	Partial Sum of Squares	Degrees of freedom	Ratio of Partial to Total Sum of Squares
Month	1015.7	11	0.305
Carrier	219.9	5	0.066
Total	3335.2	803	1.000
R-squared	0.370		

TABLE VI: MODELS FOR PREDICTING EXPECTED QUALITY.
PANEL A: DEPENDENT VARIABLE IS PERCENTAGE OF DELAYED FLIGHTS

	(1)	(2)	(3)
Month=February	-3.601	-3.601	-3.770
	(1.031)**	(1.021)**	(1.021)**
Month=March	-4.927	-4.927	-4.899
	(0.989)**	(0.978)**	(0.984)**
Month=April	-8.739	-8.766	-8.734
	(0.974)**	(0.958)**	(0.955)**
Month=May	-8.397	-8.397	-8.402
	(1.050)**	(1.033)**	(1.024)**
Month=June	-3.163	-3.163	-3.114
	(1.151)**	(1.136)**	(1.131)**
Month=July	-5.139	-5.139	-5.020
	(1.117)**	(1.097)**	(1.093)**
Month=August	-6.151	-6.151	-6.016
	(1.110)**	(1.093)**	(1.093)**
Month=September	-10.685	-10.685	-10.685
	(1.108)**	(1.117)**	(1.075)**
Month=October	-10.019	-10.019	-10.052
	(0.963)**	(0.950)**	(0.940)**
Month=November	-8.551	-8.551	-8.585
	(0.958)**	(0.949)**	(0.948)**
Month=December	-1.652	-1.652	-1.555
	(1.112)	(1.097)	(1.096)
Carrier=Continental		0.413	1.127
		(0.583)	(0.603)+
Carrier=Delta		0.944	1.073
		(0.600)	(0.608)+
Carrier=Northwest		-1.659	-1.832
		(0.772)*	(0.724)*
Carrier=TWA		0.599	1.146
		(0.698)	(0.710)
Carrier=United		2.996	3.125
		(0.755)**	(0.762)**
Carrier=US Airways		0.441	0.519
		(0.677)	(0.677)
Strike			7.070
			(4.610)
Bankruptcy			-2.972
			(0.692)**
Observations	993	993	993
R-squared	0.23	0.26	0.28
F-test: Carrier effects		5.36	7.32

Standard errors in parentheses. + significant at 10%; * significant at 5%; ** significant at 1%

TABLE VI CONTINUED

PANEL B: DEPENDENT VARIABLE IS MISHANDLED BAGS PER 1000 PASSENGERS

	(1)	(2)	(3)
Month=February	-1.669	-1.669	-1.652
	(0.316)**	(0.302)**	(0.302)**
Month=March	-1.752	-1.752	-1.751
	(0.296)**	(0.282)**	(0.280)**
Month=April	-2.677	-2.677	-2.675
	(0.278)**	(0.266)**	(0.265)**
Month=May	-2.786	-2.786	-2.787
	(0.280)**	(0.266)**	(0.265)**
Month=June	-2.014	-2.014	-2.025
	(0.297)**	(0.280)**	(0.279)**
Month=July	-2.100	-2.100	-2.106
	(0.281)**	(0.266)**	(0.264)**
Month=August	-2.217	-2.217	-2.222
	(0.287)**	(0.269)**	(0.268)**
Month=September	-3.020	-3.020	-3.020
	(0.274)**	(0.261)**	(0.260)**
Month=October	-2.659	-2.659	-2.660
	(0.286)**	(0.272)**	(0.273)**
Month=November	-2.441	-2.441	-2.442
	(0.285)**	(0.269)**	(0.269)**
Month=December	0.275	0.275	0.266
	(0.371)	(0.353)	(0.352)
Carrier=Continental		-0.144	-0.131
		(0.154)	(0.163)
Carrier=Delta		0.073	0.057
		(0.152)	(0.152)
Carrier=Northwest		0.585	0.599
		(0.145)**	(0.144)**
Carrier=TWA		1.131	1.136
		(0.197)**	(0.204)**
Carrier=United		1.088	1.072
		(0.147)**	(0.147)**
Carrier=US Airways		0.238	0.227
		(0.179)	(0.181)
Strike			-0.707
			(0.447)
Bankruptcy			-0.148
			(0.203)
Observations	1001	1001	1001
R-squared	0.30	0.37	0.37
F-test: Carrier effects		19.08	18.93

Standard errors in parentheses. + significant at 10%; * significant at 5%; ** significant at 1%

TABLE V: REGRESSIONS OF PASSENGER COMPLAINTS ON ACTUAL AND EXPECTED QUALITY
PANEL A: DEPENDENT VARIABLE IS COMPLAINTS ABOUT FLIGHT PROBLEMS PER MILLION PASSENGERS

	(1)	(2)	(3)	(4)	(5)[1]	(6)[2]
Actual delays	0.451	0.581	0.486	0.486	0.536	0.564
	(0.092)**	(0.102)**	(0.085)**	(0.094)**	(0.098)**	(0.104)**
Expected delays		-0.501	-0.470	-0.398	-0.377	-0.406
		(0.115)**	(0.101)**	(0.106)**	(0.134)**	(0.116)**
t			-0.475	-0.474		
			(0.046)**	(0.043)**		
t-squared			0.0028	0.0028		
			(0.0002)**	(0.0002)**		
Carrier=Continental				2.926		
				(1.182)*		
Carrier=Delta				-3.430		
				(0.610)**		
Carrier=Northwest				2.737		
				(0.061)**		
Carrier=TWA				4.346		
				(0.810)**		
Carrier=United				0.485		
				(0.806)		
Carrier=US Airways				-0.870		
				(0.681)		
Observations	993	993	993	993	993	993
Observations censored at 0	7	7	7	7	7	7

Results from Tobit regressions. Bootstrapped standard errors are in parentheses. Flight delays are in percent.
+ significant at 10%; * significant at 5%; ** significant at 1%
The expectation is the predicted value from column 2 in Table 4, Panel A.
[1] The expectation is the predicted value from column 1 in Table 4, Panel A.
[2] The expectation is the predicted value from column 3 in Table 4, Panel A.

TABLE V CONTINUED

PANEL B: DEPENDENT VARIABLE IS COMPLAINTS ABOUT MISHANDLED BAGGAGE PER MILLION PASSENGERS

	(1)	(2)	(3)	(4)	(5)[3]	(6)[4]
Actual mishandled bags	0.766	0.952	0.539	0.533	1.001	0.954
	(0.121)**	(0.157)**	(0.131)**	(0.116)**	(0.148)**	(0.158)**
Expected mishandled bags		-0.505	-0.108	-0.336	-0.785	-0.508
		(0.156)**	(0.143)	(0.126)**	(0.156)**	(0.151)**
t			-0.164	-0.165		
			(0.015)**	(0.015)**		
t-squared			0.0010	0.0010		
			(0.0001)**	(0.0001)**		
Carrier=Continental				1.291		
				(0.498)*		
Carrier=Delta				-1.565		
				(0.237)**		
Carrier=Northwest				-0.424		
				(0.288)		
Carrier=TWA				3.314		
				(0.482)**		
Carrier=United				-0.401		
				(0.263)		
Carrier=US Airways				-1.161		
				(0.230)**		
Observations	1001	1001	1001	1001	1001	1001
Observations censored at 0	23	23	23	23	23	23

Results from Tobit regressions. Bootstrapped standard errors are in parentheses. Mishandled bags are per thousand passengers. + significant at 10%; * significant at 5%; ** significant at 1%
The expectation is the predicted value from column 2 in Table 4, Panel B.
[3] The expectation is the predicted value from column 1 in Table 4, Panel B.
[4] The expectation is the predicted value from column 3 in Table 4, Panel B.

TABLE VI: REGRESSIONS OF PASSENGER COMPLAINTS ON THE PREDICTION ERROR
PANEL A: DEPENDENT VARIABLE IS COMPLAINTS ABOUT FLIGHT PROBLEMS PER MILLION PASSENGERS

	(1)	(2)	(3)	(4)	(5)	(6)
(Actual - expected) delays	0.581	0.486	0.486			
	(0.102)**	(0.097)**	(0.097)**			
Positive deviation from expectation				0.901	0.847	0.824
(worse than expected)				(0.174)**	(0.152)**	(0.166)**
Negative deviation from expectation				0.032	-0.142	-0.102
(better than expected)				(0.114)	(0.100)	(0.100)
t		-0.475	-0.475		-0.482	-0.482
		(0.050)**	(0.050)**		(0.047)**	(0.049)**
t-squared		0.0028	0.0028		0.0029	0.0029
		(0.0003)**	(0.0003)**		(0.0003)**	(0.0003)**
Carrier=Continental			2.963			3.407
			(1.307)*			(1.374)**
Carrier=Delta			-3.347			-2.914
			(0.547)**			(0.632)**
Carrier=Northwest			2.593			2.306
			(0.904)**			(0.882)**
Carrier=TWA			4.399			4.166
			(0.707)**			(0.883)**
Carrier=United			0.749			0.487
			(0.745)			(0.629)
Carrier=US Airways			-0.832			-0.951
			(0.607)			(0.676)
Observations	993	993	993	993	993	993
Observations censored at 0	7	7	7	7	7	7

Results from Tobit regressions. Bootstrapped standard errors are in parentheses. Flight delays are in percent.
+ significant at 10%; * significant at 5%; ** significant at 1%

TABLE VI CONTINUED

PANEL B: DEPENDENT VARIABLE IS COMPLAINTS ABOUT MISHANDLED BAGGAGE PER MILLION PASSENGERS

	(1)	(2)	(3)	(4)	(5)	(6)
(Actual - expected) mishandled bags	0.952	0.531	0.529			
	(0.166)**	(0.129)**	(0.115)**			
Positive deviation from expectation				1.606	0.956	0.742
(worse than expected)				(0.335)**	(0.278)**	(0.258)**
Negative deviation from expectation				-0.173	-0.144	0.193
(better than expected)				(0.194)	(0.147)	(0.158)
t		-0.165	-0.165		-0.156	-0.161
		(0.017)**	(0.014)**		(0.016)**	(0.014)**
t-squared		0.0010	0.0010		0.0010	0.0010
		(0.0001)**	(0.0001)**		(0.0001)**	(0.0001)**
Carrier=Continental			1.261			1.191
			(0.444)**			(0.444)**
Carrier=Delta			-1.551			-1.593
			(0.224)**			(0.250)**
Carrier=Northwest			-0.309			-0.327
			(0.286)			(0.290)
Carrier=TWA			3.537			3.365
			(0.457)**			(0.418)**
Carrier=United			-0.187			-0.192
			(0.220)			(0.222)
Carrier=US Airways			-1.114			-1.207
			(0.229)**			(0.236)**
Observations	1001	1001	1001	1001	1001	1001
Observations censored at 0	23	23	23	23	23	23

Results from Tobit regressions. Bootstrapped standard errors are in parentheses. Mishandled bags are per thousand passengers. + significant at 10%; * significant at 5%; ** significant at 1%

[1] Resnick *et al.* [2006] and Bajari and Hortacsu [2004] provide a list of existing studies and summarize their findings.

[2] See Resnick *et al.* and Bajari and Hortacsu.

[3] A similar behavior is predicted by social expectancy theory for human interactions [Darley and Fazio, 1980, and Zebrowitz, 1997]. This theory predicts that if individual *A* has different expectations for how individuals *B* and *C* will behave, then *A* will have different reactions to the same observed behavior by *B* and *C*.

[4] On eBay, sellers may take on a new identity and reappear under that after they have received negative feedback on a previous identity.

[5] Performance data are also available for a third complaint category, ticket oversales. However, those data are only reported on a quarterly basis and for a much shorter time series. I therefore do not include that category in the analysis.

[6] Data are also reported for America West and Southwest Airlines. I do not include those airlines because they have marketed themselves as 'low-cost' airlines and passengers' expectations of service quality may therefore be quite different for these airlines than for the other major carriers.

[7] Complaints are still only reported at the airline-month level.

[8] The estimation will take this into account by using Tobit regressions.