Independent Review and Analysis of the Impact of the Three-Hour Tarmac Delay Rule

Submitted To:

U.S. Department of Transportation
1200 New Jersey Avenue, SE
West Wing
Washington, DC 20590-0001

Submitted By:

Econometrica, Inc.
7475 Wisconsin Avenue
Suite 1000
Bethesda, MD 20814

January 9, 2014

ECONOMETRICA, INC.

Table of Contents

Executive Summary ..... **iv**

Introduction and Background **1**

Scope of Work **2**

Recent Flight Cancellation and Delay Trends **2**
Flight Cancellations and Taxi-Out Delays 3
Flights Experiencing a Taxi-Out Delay of More Than 2 Hours 4

Previous Studies of TDR Impact **5**

Data and Methodological Issues With Analyzing TDR Impact **7**

Statistical Relationship Between Cancellations and Taxi-Out Delays **9**

New Estimates of the Impact of the TDR Flight Cancellations **11**
Two Alternative Approaches to Isolate TDR Impact on Flight Cancellations 12
Calculating Baseline and Excess (TDR-Attributable) Cancellations 12
Approach 1: TDR Impact Measured as the Number of Excess Cancellations on Days With More Than One Taxi-Out Delay of More Than 2 Hours 13
Approach 2: TDR Impact Measured as Excess Cancellations on High Delay Days Defined Independently for Each Summer Season 14
Discussion of Results 16
Comparison With Other Estimates of Flight Cancellation Increases 16

Impact on Lengthy Taxi-Out Delays of Less Than 3 Hours **17**
Reductions in Taxi-Out Delays of More Than 60 Minutes 17
Reductions in Taxi-Out Delays Attributable to the TDR 17

Analysis of the Effect of the TDR on Passenger Welfare **18**

Conclusion **21**

Appendix A: Comparison of TDR Study Impact Estimates **A-1**

Appendix B: Excess Cancellations in June, July, and August **B-1**
Approach 1: TDR Impact Measured as the Number of Excess Cancellations on Days With More Than One Taxi-Out Delay of More Than 2 Hours B-1
Approach 2: TDR Impact Measured as Excess Cancellations on High Delay Days Defined Independently for Each Summer Season B-2

List of Tables

Table 1: May–September Flight Cancellations and Flights With >2-Hour Taxi-Out Delays 3
Table 2: Pre-Rule and Post-Rule Flight Cancellations and >2-Hour Taxi-Out Delays by Month . 4
Table 3: May–September Taxi-Out Delay Incidents Lasting >2 Hours 4
Table 4: Pre-Rule and Post-Rule Outcomes of Taxi-Out Delay Incidents Lasting >2 Hours 5
Table 5: Carrier-Specific Correlations Between Daily Cancellations and >2-Hour Delays 10
Table 6: Summer Departures and Cancellations by Delay Group (Approach 1) 13
Table 7: Summer Cancellations by Delay Day Group vs. 2009 Baseline (Approach 1)..... 14
Table 8: Number of >2-Hour Delays Included in Each Delay Day Group (Approach 2)..... 15
Table 9: Summer Departures and Cancellations by Delay Group (Approach 2) 15
Table 10: Summer Cancellations by Delay Day Group vs. 2009 Baseline (Approach 2)..... 16
Table 11: Number of Summer Flights With Lengthy Taxi-Out Times by Delay Interval 17
Table 12: Reductions in Summer Season >60-Minute Taxi-Out Delays vs. 2009 Baseline 18
Table 13: Cost to Passengers of Each Additional Minute of Travel Time by Delay Interval 20
Table A-1: Tarmac Delay Rule (TDR) Study Methodology, Data, and Estimated Impact on Cancellations A-1
Table B-1: June–August Departures and Cancellations by Delay Group (Approach 1) B-1
Table B-2: June–August Cancellations by Delay Day Group vs. 2009 Baseline (Approach 1) B-2
Table B-3: Number of >2-Hour Delays Included in Each June–August Delay Day Group (Approach 2) B-2
Table B-4: June–August Departures and Cancellations by Delay Group (Approach 2) B-3
Table B-5: June–August Cancellations by Delay Day Group vs. 2009 Baseline (Approach 2) B-3

Executive Summary

The Department of Transportation's Final Rule, Enhancing Airline Passenger Protections,[1] includes a section with requirements intended to minimize passenger exposure to tarmac delays of more than 3 hours ("the tarmac delay rule," or TDR). Since the TDR took effect, several studies—including the Regulatory Impact Analysis (RIA) prepared by DOT contractor HDR Engineering, Inc. (HDR) to accompany publication of the rule,[2] studies by American Aviation Institute (AAI),[3] and a Government Accountability Office (GAO) report[4]—have presented disparate assessments of the impact of the TDR:[5]

- HDR estimated that very few flights—41 in the first year—would be cancelled as a result of the TDR.
- In contrast, the AAI studies found that the TDR has caused massive increases in the numbers of flight cancellations, as carriers responded conservatively to situations with the potential for lengthy tarmac delays that may result in substantial fines, leading to adverse impact for passenger welfare.
- The GAO found that the rule is correlated with a greater likelihood of flight cancellations, leading to long overall passenger travel times.

DOT retained Econometrica to conduct an independent review of the various reports that have studied how the TDR has affected airline travel. This report analyzes the issues associated with flight cancellations and the effect on passenger welfare attributable to the TDR and presents new data and estimates of TDR impact. Our research indicates that the conclusions reached about the impact of the TDR depend on the data and the methodology used to analyze the effects of the rule and how the results of the analysis are interpreted.

Important observations and results presented in this report include the following:

- Carrier decisions to cancel specific flights are subject to many variables beyond weather conditions, airport congestion, and load factors. For example, a carrier may need to cancel flights as a result of late arriving flights or aircraft with maintenance issues, or it may do so as a result of schedule consolidations following mergers.[6] These cancellations

[1] Department of Transportation, "Enhancing Airline Passenger Protections," FR 74:249, December 30, 2009. A subsequent rulemaking expanded the requirements for carrier tarmac delay contingency plans and extended coverage of the TDR to include a 4-hour limit on tarmac delays for international flights arriving at or departing from the United States. These requirements became effective on August 23, 2011. See Department of Transportation, "Enhancing Airline Passenger Protections," FR 76:79, April 25, 2011.

[2] DOT, "Final Regulatory Impact Analysis of Rulemaking on Enhanced Airline Passenger Protections," December 17, 2009.

[3] Joshua Marks and Darryl Jenkins, "Impact of Three-Hour Tarmac Delay Rules and Fines on Passenger Travel Time and Welfare," July 20, 2010, and Joshua Marks and Darryl Jenkins, "Summer 2010 Cancellations and Five-Month Impact of Three-Hour Tarmac Delay Rule," November 18, 2010.

[4] Government Accountability Office, "Airline Passenger Protections: More Data and Analysis Needed to Understand Effects of Flight Delays," GAO-11-733, September 2011.

[5] A summary table comparing these studies is provided in Appendix A.

[6] It is unlikely that the significance of these factors in carrier cancellation decisions remained constant from year to year—consolidation of merging carrier schedules or changes in aircraft fleet composition or maintenance could have

may reduce the taxi-out delays—and thus the probability of cancellations if compliance with the TDR is a concern—not only for that carrier's flights, but also for other carriers' flights departing from the same airport. These cross-carrier effects on flight cancellations and delays suggest that the impact of the TDR may be more appropriately measured at the carrier, airport, or system level, rather than on a flight-specific basis.

- The cancellation *rate* for flights experiencing lengthy taxi-out delays has risen in the post-Rule period. However, this is primarily attributable to fewer flights incurring lengthy taxi-out delays (i.e., the denominator of the cancellation rate) than in the pre-Rule period, rather than an increase in the number of flights returning to the gate and then being cancelled (the numerator of the cancellation rate). Post-delay cancellations continued to account for less than 1 percent of total cancellations.
- There was a statistically meaningful association between the daily numbers of lengthy taxi-out delays and the total number of flight cancellations *in the pre-Rule period.* This relationship became stronger in the post-Rule period. There was a positive relationship between flight cancellations and lengthy taxi-out delays in the only pre-Rule summer (2009) for which complete OTP data were reported. Moreover, the positive correlation between these two variables was larger in each of the first two post-Rule summers (2010 and 2011) than in the summer of 2009.
- A valid measure of TDR impact should take into account changes in the baseline level of cancellations observed on the days with very few or no lengthy tarmac delays. Increases in the numbers of cancellations on these "Low Delay Days" cannot be reasonably attributed to the TDR.
- Our estimates of excess cancellations on High Delay Days in summer 2010 (491 to 2,444 flights) are larger than either the 41 additional annual cancellations projected in the initial regulatory evaluation conducted by HDR or the estimate of no excess cancellations in our 2010 preliminary analysis.[7] However, these estimates are smaller than the projected number of cancellations in the second M&J study (5,068) and the subsequent AAI estimate (8,114) derived using the GAO logistic model results.
- The near elimination of taxi-out delays of more than 3 hours since the TDR took effect was noted in the GAO report. However, reduced tarmac waiting times for 3,500 flights with taxi-out delays of more than an hour in summer 2010, 2,500 flights in summer 2011, and 4,300 flights in summer 2012 can also be reasonably attributed to the TDR.

resulted in a larger or smaller number of cancellations. However, the impact of any changes in these factors that took place between the summer of 2009 and the summer of 2010 would be attributed to the TDR in the GAO model.

[7] Our 2010 analysis was limited to the impact of the TDR on cancellations of flights that had already experienced taxi-out delays of greater than 2 hours. The HDR regulatory evaluation estimated the projected impact of the TDR only on cancellations of flights with tarmac delays in excess of 3 hours.

Based on the analysis and results presented in this report, the TDR has had some adverse and some beneficial impacts on passenger welfare during the period analyzed**.** The TDR appears to have had an adverse impact on the number of flight cancellations in the summer months of 2011. The TDR also appears to have had a smaller adverse impact on flight cancellations in 2010 and 2012 using one of the two methods developed for this analysis, although the impact of the TDR on cancellations during these two summers appears to have been minimal or non-existent using the second analytical approach presented in this report. In addition, the TDR appears to have reduced taxi-out waiting times for several thousand flights in each of the three post-Rule summers.

Introduction and Background

Lengthy tarmac delays—extended periods of time where passengers remain onboard an aircraft that is not airborne—became a source of attention after a series of well-publicized incidents led to an investigation and report by the Inspector General of the Office of the Secretary of Transportation (OST). There are three distinct types of tarmac delays, all of which are addressed in the report:

- Taxi-out delays occur when an aircraft pushes back from the gate but is not able to get clearance to depart and remains on the runway for an extended period of time. Most lengthy tarmac delays occur at taxi-out.
- Taxi-in delays occur when an aircraft has landed at its destination but a gate is not available to allow passengers to deplane. These situations are comparatively rare.
- Diversion delays occur when an aircraft cannot land at its scheduled destination and is diverted to another airport to land. The diversion airports may lack the proper facilities for handling the incoming passengers—the carrier may have no regular operations at the airport, gate availability may be limited by severe weather conditions or other flight delays, and Customs and Border Protection (CBP) staff may not be available to process travelers on diverted international flights. Lengthy delays resulting from diversion are far less common than extended taxi-out delays, although they account for a disproportionate share of the longest delays.

Concerns about inadequate carrier and airport response to address these situations prompted the development and publication of a Department of Transportation (DOT) Final Rule, "Enhancing Airline Passenger Protections," which became effective on April 29, 2010.[8] Among other things, it protects airline passengers from lengthy tarmac delays and prohibits U.S. airlines operating domestic flights from permitting an aircraft to remain on the tarmac for more than 3 hours without providing passengers with the opportunity to deplane (hereafter, the "tarmac delay rule" or the "TDR"). The accompanying Regulatory Impact Analysis (RIA) prepared by DOT contractor HDR Engineering, Inc. (HDR) estimated that very few flights—41 in the first year—would be cancelled as a result of the TDR.

However, some analysts have found that the TDR has been responsible for massive increases in the numbers of flight cancellations, as carriers respond conservatively to situations with the potential for lengthy tarmac delays that may result in substantial fines.[9] Concern about the impact on passenger welfare prompted a study of the issue by the Government Accountability Office (GAO).[10] The September 2011 GAO study assessed the TDR's impact on flight

[8] Department of Transportation, "Enhancing Airline Passenger Protections," FR 74:249, December 30, 2009. As noted above, a subsequent rulemaking expanded the requirements for carrier tarmac delay contingency plans and extended the coverage of the TDR to include a 4-hour limit on tarmac delays for international flights arriving at or departing from the United States. These requirements became effective on August 23, 2011. See Department of Transportation, "Enhancing Airline Passenger Protections," FR 76:79, April 25, 2011.

[9] The TDR allows DOT to assess a fine of up to $27,500 *per passenger* on any aircraft that has been permitted to remain on the tarmac for more than 3 hours for domestic flights or 4 hours for international flights.

[10] Government Accountability Office, "Airline Passenger Protections: More Data and Analysis Needed to Understand Effects of Flight Delays," GAO-11-733, September 2011.

cancellation rates using on-time performance (OTP) data collected by the Bureau of Transportation Statistics (BTS) during the first 5 months after the TDR took effect. The GAO analysis interpreted the results as demonstrating that the TDR has resulted in statistically significant increases in cancellation rates both for flights that have already pushed back from the gate and subsequently returned without departing and for scheduled flights that were never boarded ("proactive cancellations"). Using the GAO estimates, Darryl Jenkins and Joshua Marks of the American Aviation Institute (AAI) projected that the TDR resulted in 8,114 additional flight cancellations during the 5-month summer period between May and September 2011.

Scope of Work

DOT retained Econometrica to review and analyze data that could be used to determine the impact of the TDR on flight cancellations and delays. Particular attention was directed to assessing the strength of the results obtained in the GAO study, but our review also encompassed previous studies by HDR, the DOT contractor that performed the regulatory evaluation of the TDR; Marks and Jenkins; and a preliminary analysis prepared by our firm in September 2010. For this study, Econometrica developed new estimates of the impact of the TDR on flight cancellations and delays utilizing BTS data on flight delays and cancellations for the 5-month period from May through September in each of the pre-Rule summers (2010 through 2012).

DOT also requested that Econometrica conduct an independent study of the impact of TDR-attributable changes in flight delays and cancellations on passenger welfare. Work on this task involved research and analysis to determine the impact of delays and cancellations on both the length and discomfort associated with air travel, as well as the likely preferences of various groups of passengers among various options for possible carrier responses to handle flight schedule disruptions. We also present the results of relevant research on the valuation of travel time and discomfort that could potentially be used to monetize the TDR-attributable improvements and impairments to passenger welfare.

This report focuses on lengthy taxi-out delays, which are more numerous and more likely to result in flight cancellations (especially proactive cancellations) than extended taxi-in or diversion delays. The report presents detailed results for the 5-month May to September time period used in the GAO study.[11]

Recent Flight Cancellation and Delay Trends

U.S. air carriers that have at least 1 percent of total domestic scheduled-service passenger revenues are required to report flight-specific OTP data each month to BTS. The data include all scheduled-service flights within the United States (including territories).[12] While OTP data are available since January 1995, carriers were not required until October 2008 to report taxi-out times for flights that returned to the gate and were subsequently cancelled or reboarded.

[11] June, July, and August are the 3 months during which the largest numbers of lengthy taxi-out delays occur. Results using this more narrowly defined 3-month summer period from June through August are presented in Appendix B.

[12] A summary and description of the data available in the OTP reports are provided on the BTS Web site at http://www.bts.gov/xml/ontimesummarystatistics/src/index.xml.

Consequently, no cancellations of flights that incurred tarmac delays of greater than 2 hours can be identified in pre-October 2008 OTP data. Moreover, these earlier data do not have pre-gate return delay information reported for flights that subsequently departed.

Flight Cancellations and Taxi-Out Delays

Table 1 presents data on flight cancellations and taxi-out delays of more than 2 hours from May through September during pre-Rule (October 2008 through April 2010) and post-Rule (May 2010 through September 2012) periods.

Table 1: May–September Flight Cancellations and Flights With >2-Hour Taxi-Out Delays

Year/ Month	Scheduled Departures	Total Flights Cancelled		All Flights With >2-Hour Taxi-Out Delays		Cancelled After >2-Hour Taxi-Out Delay	
		Number	Percent	Number	Percent	Number	Percent
2009	2,763,713	28,464	1.0%	3,196	0.12%	219	0.008%
2010	2,760,546	33,532	1.2%	1,658	0.06%	225	0.008%
2011	2,625,329	46,989	1.8%	2,280	0.09%	341	0.013%
2012	2,621,479	29,027	1.1%	1,464	0.06%	193	0.007%

Source: Econometrica tabulations of BTS on-time performance data.

As Table 1 shows, the total number of May–September flight cancellations increased significantly from 2009 to 2010 and again from 2010 to 2011 before falling to nearly the 2009 level in the summer of 2012. The numbers of flights cancelled after a taxi-out delay of more than 2 hours were similar in 2009, 2010, and 2012, while there were more such cancellations in 2011. However, these post-delay cancellations continued to account for less than 1 percent of total cancellations.

Table 2 presents average monthly cancellation and taxi-out delay rates for the pre-Rule period from October 2008 through April 2010 and the post-Rule period from May 2010 through September 2012.

Table 2: Pre-Rule and Post-Rule Flight Cancellations and >2-Hour Taxi-Out Delays by Month

	Total Flights Cancelled		All Flights With >2-Hour Taxi-Out Delays		Flights Cancelled After >2-Hour Tarmac Delay	
Month	**Pre-Rule**	**Post-Rule**	**Pre-Rule**	**Post-Rule**	**Pre-Rule**	**Post-Rule**
January	2.4%	2.7%	0.07%	0.04%	0.003%	0.005%
February	3.3%	2.9%	0.07%	0.03%	0.003%	0.002%
March	1.8%	1.2%	0.06%	0.02%	0.006%	0.002%
April	1.1%	1.5%	0.06%	0.03%	0.005%	0.004%
May	*0.9%*	1.4%	*0.06%*	0.06%	*0.004%*	0.008%
June	*1.5%*	1.5%	*0.22%*	0.08%	*0.017%*	0.012%
July	*1.2%*	1.5%	*0.15%*	0.09%	*0.013%*	0.013%
August	*1.0%*	1.6%	*0.12%*	0.07%	*0.004%*	0.009%
September	*0.6%*	0.8%	*0.02%*	0.04%	*0.000%*	0.004%
October	0.8%	0.9%	0.03%	0.02%	0.001%	0.001%
November	0.7%	0.7%	0.02%	0.02%	0.000%	0.001%
December	3.0%	2.2%	0.11%	0.03%	0.010%	0.002%

Note: Cancellation rate is average of 2 or 3 years except where shown in *italics*.
Source: Econometrica tabulations of BTS on-time performance data.

Table 2 shows that the average cancellation rates for September, October, and November have been lower than in any of the other 9 months in both the pre-Rule and post-Rule periods. However, the "summer season" results presented in this report incorporate data for September to ensure comparability with the analysis presented in the GAO report.

Flights Experiencing a Taxi-Out Delay of More Than 2 Hours

Table 3 shows that there were substantially fewer flights with taxi-out delays in each of the three post-Rule summers than in summer 2009, but that a higher percentage of these delayed flights were subsequently cancelled in each post-Rule summer than in summer 2009.

Table 3: May–September Taxi-Out Delay Incidents Lasting >2 Hours

		Departed After Taxi-Out Delay		Returned to Gate and Reboarded		Cancelled After Tarmac Delay		Percent of Delayed Flights Cancelled
Year	**Taxi-Out Delays**	**Number**	**Percent**	**Number**	**Percent**	**Number**	**Percent**	
2009	3,196	2,584	81%	393	12%	219	7%	6.9%
2010	1,658	1,041	63%	392	24%	225	14%	13.6%
2011	2,280	1,335	59%	604	26%	341	15%	15.0%
2012	1,464	877	60%	394	27%	193	13%	13.2%

Source: Econometrica tabulations of BTS on-time performance data.

While larger *percentages* of flights experiencing taxi-out delays of more than 2 hours were ultimately cancelled after returning to the gate in all three post-Rule summers than was the case

in 2009, the total number of flights cancelled after a gate return in summer 2012 was actually lower than in 2009.

Table 4 presents the average percentages of delayed flights in each month in the pre-Rule and post-Rule period that resulted in the following:

- Departed without returning to the gate
- Returned to the gate and are reboarded
- Were cancelled after returning to the gate

Table 4: Pre-Rule and Post-Rule Outcomes of Taxi-Out Delay Incidents Lasting >2 Hours

	Departed After Taxi-Out Delay		Returned to Gate and Reboarded		Cancelled After Tarmac Delay	
Month	**Pre-Rule**	**Post-Rule**	**Pre-Rule**	**Post-Rule**	**Pre-Rule**	**Post-Rule**
January	81%	77%	14%	13%	5%	10%
February	88%	66%	8%	27%	5%	7%
March	79%	70%	10%	22%	11%	9%
April	77%	75%	15%	15%	8%	10%
May	*87%*	60%	*7%*	25%	*6%*	15%
June	*78%*	59%	*14%*	25%	*8%*	16%
July	*80%*	60%	*12%*	25%	*9%*	15%
August	*84%*	59%	*12%*	28%	*4%*	13%
September	*82%*	67%	*16%*	24%	*2%*	9%
October	89%	72%	9%	21%	3%	6%
November	86%	73%	13%	21%	1%	6%
December	80%	75%	11%	19%	9%	6%

Note: Cancellation rate is average of 2 or 3 years except where shown in italics.
Source: Econometrica tabulations of BTS on-time performance data.

Table 4 shows that a lower percentage of flights with taxi-out delays of more than 2 hours departed without returning to the gate in May, June, July, and August than in any of the other 8 months in the post-Rule period.

The BTS data presented in this section indicate that the cancellation and delay rates for May, June, July, and August are similar to each other and unlike those for the other 8 months. However, as noted above, the results presented in this report incorporate data for September to ensure comparability with the analysis presented in the GAO report.

Previous Studies of TDR Impact

The GAO report is one of several studies that have addressed the impact of the TDR during the first 2 years after it took effect on April 29, 2010: [13]

[13] As noted above, a summary table comparing these studies is provided in Appendix A.

- In a July 20, 2010, report titled "Impact of Three-Hour Tarmac Delay Rules and Fines on Passenger Travel Time and Welfare," aviation industry consultants Joshua Marks and Darryl Jenkins (hereafter, "M&J") concluded that the tarmac delay time limit had resulted in a substantially larger number of cancellations than projected in the DOT RIA. Based on an analysis of BTS tarmac delay and flight cancellation data for May 2010, the authors projected an annual increase of 2,600 flight cancellations from carrier operations required to avoid the risk of a flight exceeding the 3-hour limit on tarmac delays. The M&J report also estimated that an additional 2,600 "follow-on" cancellations would result from lack of aircraft availability at the destination airports.[14]

- A September 2010 Econometrica report also presented an analysis of BTS tarmac delay and flight cancellation data for May, June, and July 2010, which indicated that no additional cancellations during those 3 months could be attributed to the TDR.[15] Based on these results and subsequent in-house updates and refinements to the analysis, the DOT issued press releases and other statements reiterating its position that the tarmac delay rule had not increased flight cancellations.[16]

- Marks and Jenkins subsequently updated and extended their initial critique of the TDR. The estimates in their November 2010 report incorporated data for the first 5 post-TDR months (May–September 2010). Their analysis indicated that system-wide cancellations were up 18 percent over the previous year, despite more favorable weather conditions. Almost 95 percent of the projected increase was accounted for by flights that were cancelled before leaving the gate, and 30 percent of the projected 5,068 additional cancellations involved flights at two large hub airports (O'Hare and Detroit Metro).[17]

- The second M&J report estimated that approximately 150,000 passengers had been adversely affected by the need to begin returning to the terminal at the 2.5-hour mark to ensure arrival at the gate within 3 hours. However, the bulk of the impact projected in this analysis is attributable to the impact of pre-departure cancellations, including flights for which no aircraft had been assigned. These are termed "proactive" cancellations, although there is no specific information available to indicate whether the possibility of a lengthy tarmac delay was considered in making the decision to cancel these flights.

- In March 2011, DOT provided a more detailed analysis that supported the conclusions that were reported in the October 12, 2010, press release. The DOT methodology differed in certain respects from that used in the September 2010 Econometrica report. For example, the DOT calculations excluded cancellations that were related to security or due to the air carrier. Unlike the Econometrica analysis, the DOT cancellation rates estimate proactive cancellations, which were identified as those scheduled to arrive and depart from airports with more than one 2-hour tarmac delay on the same date. The DOT analysis indicated that the total number of tarmac delay-attributable cancellations

[14] Joshua Marks and Darryl Jenkins, "Impact of Three-Hour Tarmac Delay Rules and Fines on Passenger Travel Time and Welfare," July 20, 2010.
[15] Econometrica, Inc., "Review and Analysis of the Marks/Jenkins Report on the Impact of the 3-Hour Tarmac Delay Provision of the Enhanced Airline Passenger Protections TDR," September 16, 2010.
[16] See, for example, Department of Transportation, "Long Tarmac Delays in August Show Steep Drop from Last Year," DOT 186-10, October 12, 2010.
[17] Joshua Marks and Darryl Jenkins, "Summer 2010 Cancellations and the Five-Month Impact of the Three-Hour Tarmac Rule," November 18, 2010.

declined substantially from 8,696 during the pre-TDR period from May through October 2009 to 7,120 during the first 5 full months that the TDR was in effect.[18]

- The third M&J report released in March 2011 reiterated the points made in the previous reports, amplifying the previous focus on the overall flight cancellation rate as a barometer of TDR impact. The authors took issue with the DOT (and previous Econometrica) analysis based on examining trends in flights that experience at least a 2-hour tarmac delay, arguing that the threat of "extreme fines" significantly reduced the number of such flights after the TDR took effect. However, while they do not agree that the number of flights with longer-than-2-hour tarmac delays is a valid indicator of reduced waiting times, their updated methodology uses this metric to compute the tarmac delay cancellation rate.[19]
- In September 2011, GAO released an analysis of the TDR's impact that follows the M&J studies' approach in several respects. The GAO study adopts the same metric (flight cancellation rates, rather than numbers of flights cancelled) to assess the TDR's impact and analyzes OTP data from the 5-month post-TDR period (May–September 2010) used in the second M&J study. The GAO report used a logistic regression model to characterize the statistical relationship between cancellation rates for flights in four tarmac delay interval lengths (pre-departure, 0–1, 1–2, 2 or more hours) and various potential explanatory variables. The results indicate that statistically significant increases in cancellations for flights took place from 2009 to 2010 in all four delay intervals analyzed.[20]
- Marks and Jenkins' AAI released an analysis of airline passenger protection regulations in November 2011, which includes updated estimates of TDR impact of flight cancellations based on the GAO study model (8,114 in 2010 and 13,087 in 2011).[21]

Data and Methodological Issues With Analyzing TDR Impact

There are some unavoidable data limitations that constrain efforts to examine and analyze trends in flight cancellations and delays in recent years:

- Data on flights that returned to the gate after a lengthy tarmac delay and were subsequently cancelled or reboarded are not available before October 2008. Moreover, the 19-month period from October 2008 through April 2010 includes only a single May–

[18] Department of Transportation, "Three-Hour Tarmac Delay Rule Does Not Increase Cancellations," March 14, 2011.
[19] These findings are summarized in a presentation, executive summary, and set of exhibits. See Joshua Marks and Darryl Jenkins, "The Tarmac TDR: Adjustments Needed," March 29, 2011 (Webinar).
[20] Government Accountability Office, *op. cit.*, 2011.
[21] Jenkins, Marks, and Miller, "Consumer Regulation and Taxation of the U.S. Airline Industry Estimating the Burden for Airlines and the Local Impact," November 2011.

September summer season.[22] This means that comparable OTP data are available for only a single pre-Rule summer.

- The group of U.S. carriers required to report OTP data has changed in each post-Rule summer. Two regional carriers (Comair and Pinnacle) ceased reporting in January 2011, and another (Mesa) was no longer required to report as of January 2012.[23] This means that the profiles of flights for which cancellation and delay data are reported in each year are not directly comparable to those for the other post-Rule summers.

There are also some methodological issues that need to be addressed in any analysis of TDR impact:

- The GAO study used the percentage of flights cancelled to characterize the impact of the TDR on flight cancellations. One concern is that the cancellation *rate* for flights experiencing lengthy taxi-out delays has risen in the post-Rule period is not primarily because the number of flights that returned to the gate and were subsequently cancelled (the numerator of the cancellation rate) increased, but instead because fewer flights incurred lengthy taxi-out delays (the denominator of this cancellation rate). The reduction in the number of delayed flights unambiguously represents a welfare improvement because fewer passengers are delayed. In our view, estimated changes in the *numbers* of TDR-attributable flight cancellations—and the numbers of passengers on those flights—represent a better measure of the impact of the TDR on passenger welfare than do changes in cancellation rates.

- Interpreting the results from the GAO logistic regression models depends critically on whether or not changes in the baseline level of flight cancellations observed from 2009 to 2010, after taking into account variations in weather, air traffic patterns, and other factors, are properly regarded as being attributable to the TDR. One concern is that the number of flight cancellations rose substantially in the summer of 2010 and again in 2011, relative to 2009, *even on days when carrier compliance with the TDR was unlikely to have been a concern.* In our view, estimates of TDR on flight cancellations should control for changes in the baseline level of cancellations that took place on days in each of post-Rule summer when few, if any, lengthy taxi-out delays occurred.

- The GAO study does not includes an assessment of TDR impact on taxi-out delays of less than the 3 hours, the threshold after which passengers must be offered the option to deplane. We include an assessment of TDR impact on these delays in our analysis.

[22] A substantial majority of lengthy taxi-out delays take place during this time period, typically as a result of sudden or unpredictable onsets of severe weather events. Consequently, analysis and estimation of TDR-attributable cancellations has focused on the summer months.

[23] There were also reporting consolidations during this time period as a result of the Air Tran–Southwest, Continental–United, Delta–Northwest, and Express Jet–Sky West mergers. Overall, there were 19 reporting carriers in 2009 and only 12 remaining in 2012.

Statistical Relationship Between Cancellations and Taxi-Out Delays

The GAO logistic regression models take into account pre-Rule and post-Rule variation in weather conditions, airport congestion, load factors, and several other variables that may influence the probability that an individual flight could be cancelled. These models yield estimated cancellation probabilities for individual flights, which are then used to characterize the potential impact of the TDR.[24]

However, the narratives in the GAO report, M&J studies, and individual carrier comments, all describe decision-making processes that take into account the impact of prospective flight cancellations on a carrier's entire daily operations. The GAO approach does not seem to be particularly well-suited to capture these cross-flight, cross-carrier, or system-wide interaction effects. For example, a carrier may need to increase flight cancellations on a given day in response to late arriving flights or aircraft with maintenance issues.[25] These cancellations may reduce the risk of lengthy taxi-out delays—and thus the probability of cancellations if compliance with the TDR is a concern—not only for that carrier's flights, but also for other carriers' flights departing from the same airport.

The potential impact of these carrier decisions based on these cross-carrier and system-wide factors can be captured by analyzing cancellation and delay data on a daily basis, rather than for each individual flight. The simplest, and most obvious, approach to assessing the impact of the TDR, in our view, is to investigate the direct relationship between:

- The events that trigger the need to make these decisions—in most cases, the possibility of several lengthy taxi-out delays on a given day.
- The outcomes of those decisions, which are reflected in the number of additional cancellations over the baseline level that would have taken place in the absence of concerns about complying with the TDR.

Preliminary evidence on the strength of the pre-Rule and post-Rule relationships between these two variables can be derived from OTP data on daily cancellations and lengthy taxi-out delays. For our analysis, we computed the Pearson correlations coefficients for flight cancellations and taxi-out delays of more than 2 hours for each reporting carrier, each large hub airport that experiences lengthy taxi-out delays, and for the system as a whole to determine if:

- There was a statistically meaningful association between the daily numbers of lengthy taxi-out delays and flight cancellations in the pre-Rule period. If these two variables were weakly related in the pre-Rule period, the number of lengthy taxi-out delays would not be likely to serve an adequate predictor of cancellations in the post-Rule period.

[24] Separate models are estimated for gate cancellations and for cancellations of flights with taxi-out delays. See GAO Report, Appendix V, "Tarmac Delay Logistic Regression Analysis."

[25] It is possible that the influence of each of these factors in carrier cancellation decisions changes over time. For example, consolidation of merging carrier schedules or changes in aircraft fleet composition or maintenance could have resulted in a larger or smaller number of cancellations. However, the GAO approach would implicitly attribute the impact of any changes in these factors that took place between summer 2009 and summer 2010 to the TDR.

- This relationship became stronger in the post-Rule period. The penalties for non-compliance specified in the TDR should mean that carriers would be more likely to cancel flights on days in which the probability of one or more departing flights experiencing a taxi-out delay of more than 3 hours is higher.

Table 5 presents the results of a preliminary analysis we conducted using OTP data from the 3-month periods from June through August in 2009, 2010, and 2011. Correlation coefficients were calculated between the numbers of delays of more than 2 hours at large hub airports for reporting carriers and the numbers of flight cancellations on the same day. Note that a correlation coefficient of 1 reflects a uniformly positive association between the numbers of cancellations and lengthy taxi-out delays on any given day during the summer season for each carrier. A coefficient of 0 indicates that the two variables are unrelated.[26]

Table 5: Carrier-Specific Correlations Between Daily Cancellations and >2-Hour Delays

Carrier/Category	2009	2010	2011*
American	0.15	0.21	0.48
Delta	0.66	0.53	0.66
United	0.44	0.40	0.50
US Airways	0.64	0.61	0.61
Continental	0.53	0.15	0.14
Mainline Carriers (5)	**0.37**	**0.62**	**0.72**
AirTran	0.56	0.31	0.53
Frontier	0.26	0.07	-0.04
JetBlue Airways	0.60	0.16	0.48
Southwest Airlines	0.52	0.20	0.53
Low Cost (4)	**0.71**	**0.29**	**0.61**
American Eagle	0.60	0.47	0.45
Atlantic Southeast	0.41	0.63	0.46
ExpressJet	0.69	0.40	0.42
SkyWest	0.41	0.48	0.51
Mesa Airlines	0.33	0.57	0.24
Regional Carriers (5)	**0.61**	**0.55**	**0.52**
Alaska Airlines	-0.09	0.13	0.02
All Reporting Carriers (16)**	**0.51**	**0.61**	**0.69**

*Excludes August 27–29, 2011 (Hurricane Irene)

**Includes Hawaiian Airlines (HA), which had no taxi-out delays of more than 2 hours.

[26] A correlation coefficient could be not calculated for Hawaiian Airlines, which experienced no taxi-out delays in excess of 2 hours during any of the three summers. The 2009 coefficient for Delta Airlines includes cancellation and delay counts for Northwest Airlines, which consolidated reporting at the beginning of 2010. No correlations were calculated for Comair and Pinnacle, which ceased reporting OTP data at the beginning of 2011. The 2011 correlation coefficients exclude the data for the 3-day period from August 27 through August 29. There were 6,800 cancellations during these 3 days, most of which were attributable to the impact of Hurricane Irene, but there were no taxi-out delays of more than 2 hours.

These results demonstrate that the relationship between cancellations and lengthy taxi-out delays has been consistently positive in both pre-Rule and post-Rule summers. More specifically:

- For all reporting carriers, generally there was a positive relationship between cancellations and delays in all three summer seasons.
- This relationship was somewhat stronger in each of the two post-Rule summers than in 2009.
- The association between cancellation and delays was much stronger for the mainline carriers in the post-Rule summers, while the relationship between cancellation and delays was less strong for the low-cost and regional carriers in the 2010 and 2011 summers than in 2009.
- The relationship between cancellations and delays during these three summers was consistently stronger for Delta and US Airways than for any of the other mainline carriers.

Two other conclusions from this analysis should be noted:

- It is reasonable to expect that the correlation coefficients were stronger and more stable from year to year for the system as a whole than for most individual carriers, whose delay and cancellation experiences may be more influenced by specific operational factors or extreme weather events at individual hub airports.
- Year-to-year comparison of the correlation coefficients that were above the system-wide average in the previous year suggests that carriers work over time to mitigate the numbers of cancellations resulting from high numbers of taxi-out delays, not only by reducing the numbers of these delays, but also by reducing the number of flights cancelled in response to any given number of taxi-out delays that do occur. In other words, the relationship between lengthy taxi-out delays and flight cancellations may be altered over time by carrier scheduling and operations improvements that address situations that could potentially result in violations of TDR requirements.

New Estimates of the Impact of the TDR Flight Cancellations

If more flights are cancelled on days with larger numbers of taxi-out delays in excess of 2 hours, one way to assess the impact of the TDR would be to compare the increases in cancellations during the pre-Rule and post-Rule summers on days that had relatively few, if any, lengthy taxi-out delays with the increase in cancellations on days with relatively large numbers of lengthy taxi-out delays. If the TDR had an adverse impact, the post-Rule increase in cancellations on the days with *highest* numbers of lengthy taxi-out delays should have larger than the post-Rule increase in cancellations that occurred on days with the *fewest* numbers of lengthy taxi-out delays.

In other words, the post-Rule increase in cancellations on days when there were minimal numbers of these delays can be used to determine if the baseline level of flight cancellations increased from the pre-Rule period on days when compliance with the TDR was not likely to have been a concern. If there were larger post-Rule increases in the numbers of cancellations on

days with higher numbers of lengthy taxi-out delays, relative to the increases on days with few if any such delays, these above-baseline cancellations on days with higher numbers of taxi-out delays in excess of 2 hours could be fairly attributed to the TDR.

Two Alternative Approaches to Isolate TDR Impact on Flight Cancellations

For this analysis we used two alternative approaches to isolate the potential impact of the TDR on flight cancellations. The first approach divided all summer days into two categories:

- Low Delay Days were defined as those with no more than one reported taxi-out delay of more than 2 hours at any large hub airport.
- Days with more than one lengthy taxi-out delay were included in the High Delay Day group.

Under this approach there were a smaller number of days in the High Delay Day group in the post-Rule summers than in the pre-Rule summer because the number of summer days with more than one flight experiencing a taxi-out delay in excess of 2 hours fell substantially after the TDR took effect. One concern with this approach is that it implicitly assumes that the post-Rule increases in the numbers of Low Delay Days should *not* be attributed to the TDR, although the Rule may have prompted changes in carrier and airport flight operations management that contributed to these increases.

The second approach divides each summer season into three groups of days—Low Delay Days, Moderate Delay Days, and High Delay Days. In contrast to the first approach, the number of lengthy taxi-out delays that used to assign days to each of the delay day categories is allowed to vary from year to year to ensure that there are approximately equal numbers of days in each group. In other words, the Low Delay Day group includes the "best" days for each summer, while the High Delay Day group includes the "worst" days. This (second) approach allows for the possibility that the TDR contributed to the observed reductions in the numbers of lengthy taxi-out delays observed in all of the post-Rule summers.

Calculating Baseline and Excess (TDR-Attributable) Cancellations

Both of these approaches for assessing the impact of the TDR on passenger welfare involve developing estimates of the increases in flight cancellations and reductions in lengthy taxi-out delays that would have been observed in each post-Rule summer if the TDR had not been implemented:[27]

- Baseline (non TDR-attributable) increases in cancellations are computed as the increase in cancellations from the pre-Rule summer and each post-Rule summer on Low Delay Days.

[27] This is *not* the same as estimating as the baseline for comparison the number of cancellations that would have taken place in the summer of 2009, adjusting for differences in weather and air traffic characteristics between two time periods, which is the approach used in the GAO study. The relevant benchmark is the number of cancellations that **would have occurred** given the carrier, airport, and air traffic control operations environment that existed in each of the post-Rule summers.

- Excess cancellations are computed by subtracting these baseline increases from the increases in actual cancellations from the pre-Rule summer and each post-Rule summer on High Delay Days.

If the TDR had a measurable impact on flight cancellations, the observed increase in cancellations on days with larger numbers of lengthy taxi-out delays occurred should have been larger than on other days during the same time period. These excess cancellations are arguably attributable to the TDR, while increases in cancellations on days with few, if any, lengthy taxi-out delays at any U.S. airports are not likely to be related to the TDR.[28]

Approach 1: TDR Impact Measured as the Number of Excess Cancellations on Days With More Than One Taxi-Out Delay of More Than 2 Hours

The first set of excess cancellation estimates were developed using cancellations on days with either zero or one taxi-out delay of more than 2 hours as the Low Delay Day baseline. All days with more than one lengthy taxi-out delay were included in the High Delay Day group.

Table 6 shows the numbers of departures and cancellations for these two groups in each summer season. [29]

Table 6: Summer Departures and Cancellations by Delay Group (Approach 1)

Year	Delay Day Group	Scheduled Departures	Actual Cancellations	Adjusted Cancellations
2009	0/1	833,295	4,589	4,589
2009	>1	1,930,418	23,875	23,875
2010	0/1	1,232,987	10,644	7,194
2010	>1	1,527,559	22,888	28,924
2011*	0/1	988,264	9,785	8,251
2011	>1	1,588,409	30,405	36,952
2012	0/1	1,153,226	8,736	6,312
2012	>1	1,468,253	20,291	26,678

*Excludes August 27–29, 2011 (Hurricane Irene)
Source: Econometrica tabulations of BTS on-time performance data.

The "actual cancellations" numbers in Table 6 vary in part because of differences in the number of scheduled departures in the same Delay Day group during pre-Rule and post-Rule summers. The "adjusted cancellations" scale actual cancellations in each post-Rule summer to estimate the numbers of cancellations that would have taken place if the number of scheduled departures in that that delay day group was unchanged from the number in the summer of 2009. These

[28] In contrast, the GAO report implicitly attributes the entire observed increase in cancellations from 2009 to 2010 to the TDR, including the increases that occurred on days with few, if any, lengthy taxi-out delays at any U.S. airports.
[29] These calculations exclude the OTP data for the 3 days in August 2011 during which Hurricane Irene caused thousands of cancellations. Because there were no taxi-out delays of more than 2 hours, including these 3 days in the Low Delay Day group would unrealistically increase the number of baseline cancellations in 2011, resulting in large *negative* estimates of excess cancellations.

adjusted cancellation estimates can then be used to compute the number of excess cancellations on High Delay Days in each post-Rule summer. These calculations are shown in Table 7.

Table 7: Summer Cancellations by Delay Day Group vs. 2009 Baseline (Approach 1)

Number of >2-Hour Taxi-Out Delays	2010	2011	2012
0/1	2,605	3,662	1,723
>1	5,049	13,077	2,803
Excess Cancellations Over Baseline	**2,444**	**9,415**	**1,080**

*Excludes August 27–29, 2011 (Hurricane Irene)

Based on this approach, 2,444 cancellations in summer 2010, 9,415 cancellations in summer 2011, and 1,080 cancellations in 2012 can be reasonably attributed to the TDR. However, there are valid reasons to expect that this approach may not provide reliable estimates of excess cancellations:

- There were fewer than 2 lengthy taxi-out delays on only 19 of the 92 summer days in 2009, which means that the baseline cancellation estimate for Low Delay Days is based on a relatively small percentage of the available flight data for the summer of 2009.
- The increased number of Low Delay Days in post-Rule summers (37 in 2010 and 2012; 33 in 2011) is treated as being unrelated to the TDR.

The second approach addresses these concerns.

Approach 2: TDR Impact Measured as Excess Cancellations on High Delay Days Defined Independently for Each Summer Season

As noted above, one possible concern with the first approach is that there were substantially fewer summer days with more than one extended taxi-out delay in the post-Rule summers than in 2009. The days on which reductions in the numbers of the lengthy taxi-out delays actually took place are most likely to have been those that would have occurred on the days with the most flexibility in carrier and airport departure schedules. Consequently, management of flight operations on High Delay Days in the post-Rule summers would not be directly comparable to the relatively less difficult management challenges on average that carriers faced on the larger number of High Delay Days in the summer of 2009.

This problem can be addressed by dividing the 153 days in each 5-month summer season into three approximately equal groups of Low, Moderate, and High Delay Days. As with the first approach, days are assigned to groups based on the numbers of flights that experienced taxi-out delays of more than 2 hours. In contrast to the first approach, however, the minimum and maximum numbers of lengthy taxi-out delays used to define the Low, Moderate, and High Delay Day groups are not constrained to be the same in every year. Because there were substantially fewer days with large numbers of lengthy taxi-out delays in each post-Rule summer, relative to 2009, days with relatively fewer of such delays must be included in the Moderate and High Delay Day groups in the post-Rule summer seasons to ensure that the three groups contain approximately equal shares of all scheduled departures in each of the four summers. Table 8

reports the numbers of more-than-2-hour taxi-out delays per day used to define the three delay day groups in each of the four summers.

Table 8: Number of >2-Hour Delays Included in Each Delay Day Group (Approach 2)

Delay Day Group	2009	2010	2011*	2012
Low	0-1	0	0	0
Moderate	2-13	1-7	1-7	1-5
High	>13	>7	>7	>5

*Excludes August 27–29, 2011 (Hurricane Irene)

Table 9 shows the actual and adjusted numbers of departures and cancellations for the Low, Moderate, and High Delay Day groups in each pre-Rule and post-Rule summer season.

Table 9: Summer Departures and Cancellations by Delay Group (Approach 2)

Year	Delay Day Group	Scheduled Departures	Actual Cancellations	Adjusted Cancellations
2009	Low	833,295	4,589	4,589
2009	Moderate	981,060	9,387	9,387
2009	**High**	**949,358**	**14,488**	**14,488**
2010	Low	901,173	7,057	6,525
2010	Moderate	953,334	10,332	10,632
2010	**High**	**906,039**	**16,143**	**16,915**
2011*	Low	673,885	6,073	7,510
2011	Moderate	907,399	11,524	12,459
2011	**High**	**995,389**	**22,593**	**21,548**
2012	Low	963,538	7,298	6,826
2012	Moderate	833,933	7,456	8,524
2012	**High**	**824,008**	**14,273**	**15,694**

As with the first approach, the numbers of actual cancellations must be adjusted to account for variations in the numbers of scheduled departures among the pre-Rule and post-Rule summer season in each Delay Day group. However, these adjustments are much smaller numerically than those required using the first approach, because the numbers of days (and thus the number of flights) in each of the three Delay Day groups are similar in the pre-Rule and post-Rule summers.

In Table 10 these normalized post-Rule cancellations numbers are used to calculate the estimates of excess cancellations.

Table 10: Summer Cancellations by Delay Day Group vs. 2009 Baseline (Approach 2)

Delay Day Group	2010	2011	2012
Low	1,936	2,921	2,237
High	2,427	7,060	1,206
Excess Cancellations Over Baseline	**491**	**4,139**	**None**

*Excludes August 27–29, 2011 (Hurricane Irene)

Using this approach, there were 491 excess cancellations in summer 2010, 4,139 in summer 2011, and none at all in summer 2012 that can be attributed to the TDR.

Discussion of Results

The excess cancellation estimates developed using the first approach suggest that substantial numbers of TDR-attributable flight cancellations took place in each of the three post-Rule summers. It is possible that the results obtained with the first approach represent conservatively stated *upper bound* estimates of the numbers of TDR-attributable flight cancellations—just over 1,000 flights in the summer 2012 but more than 9,400 in the summer of 2011.

In contrast, the excess cancellation estimates developed using the second approach indicate that there were only a few hundred TDR-attributable flight cancellations in the summer of 2010 and none at all in the summer of 2012. However, more than 4,100 excess cancellations in the summer of 2011 can be reasonably attributed to the TDR.[30]

To test the sensitivity of both approaches to changes in the flight cancellation and delay data, we also developed estimates of baseline and excess cancellations for the 3-month period from June through August for each post-Rule year. These results are provided in Appendix B.

These estimates cannot be extrapolated to provide estimates of the 12-month impact of the TDR on flight cancellation patterns because of the extreme seasonal variation in flight cancellation and delay patterns. The past four winters have produced wildly varying numbers of cancellations, with one "good" and one "bad" winter each in the pre-Rule and post-Rule periods. There are also very few days in the fringe months of March, April, October, and November with significant numbers of lengthy taxi-out delays, which means that the excess cancellation calculations used to estimate TDR impact would be based on unreliably small samples of High Delay Days and Moderate Delay Days.

Comparison With Other Estimates of Flight Cancellation Increases

The range of estimates for excess cancellations on High Delay Days in the first post-Rule summer of 2010 (491 to 2,444 flights) from our two alternative approaches are larger than the 41 additional annual cancellations projected in the HDR regulatory evaluation or the no excess-

[30] It should be noted that the increase in cancellations on Moderate Delay Days was actually *lower* than on Low Delay Days in each of the three post-Rule summers. This result could be partially attributed to carrier, airport, and air traffic control implementation of operational improvements that significantly mitigate the risk of carrier exposure to fines for taxi-out delays of more than 3 hours on days with moderate or isolated weather events.

cancellation estimated in our 2010 preliminary analysis.[31] However, these estimates are smaller than the projected number of TDR-attributable cancellations in the second M&J study (5,068 flights in 2010) and the numbers that AAI estimated using the GAO logistic model results (8,114 flights in 2010 and 13,087 in 2011).

Impact on Lengthy Taxi-Out Delays of Less Than 3 Hours

The near-elimination of taxi-out delays of more than 3 hours since the TDR took effect was noted in the GAO report. However, substantially larger numbers of flights experienced taxi-out delays that exceeded the 60-, 90-, 120-, or 150-minute thresholds than delays lasting more than 3 hours (180 minutes) in every year before the TDR took effect. Reductions in the numbers of flights that experienced these lengthy taxi-out delays can be reasonably attributed to the TDR. However, an assessment of the impact of the TDR on taxi-out delays of 180 minutes or less is not included in the GAO report.

Reductions in Taxi-Out Delays of More Than 60 Minutes

The tabulations of the OTP data presented in Table 11 show that there have been large reductions in the numbers of summer flights with taxi-out delays of 60 minutes or more in each of the post-TDR summers, relative to 2009.

Table 11: Number of Summer Flights With Lengthy Taxi-Out Times by Delay Interval

Year	61 to 90 Minutes	91 to 120 Minutes	121 to 150 Minutes	151 to 180 Minutes	> 180 Minutes	Total (> 60 Minutes)
2009	18,297	5,303	2,012	698	484	**26,794**
2010	14,849	4,265	1,356	284	11	**20,765**
2011	12,325	4,106	1,815	433	19	**18,698**
2012	11,298	3,176	1,217	231	2	**15,924**
2010 vs. 2009	-3,448	-1,038	-656	-414	-473	**-6,029**
2011 vs. 2009	-5,972	-1,197	-197	-265	-465	**-8,096**
2012 vs. 2009	-6,999	-2,127	-795	-467	-482	**-10,870**

Source: Econometrica tabulations of BTS on-time performance data.

Passengers on 6,000 fewer summer flights in 2010, 8,000 fewer flights in 2011, and almost 11,000 fewer flights in 2012 experienced tarmac delays of more than 60 minutes length before taking off.

Reductions in Taxi-Out Delays Attributable to the TDR

Since the TDR took effect, carriers monitor departing flights that have remained on the tarmac in excess of 1 hour more closely than before. These improvements in the management of ground operations have almost certainly contributed to substantial reductions in the numbers of flights

[31] Our 2010 analysis was limited to the impact of the TDR on cancellations of flights that had already experienced taxi-out delays of greater than 2 hours. The HDR regulatory evaluation estimated the projected impact of the TDR only on cancellations of flights with tarmac delays in excess of 3 hours.

experiencing lengthy taxi-out delays of 3 hours or less, even on days where no taxi-out delays of more than 2 hours occurred.

However, the observed reductions in lengthy taxi-out flight delays may also be partly attributable to lower baseline levels of lengthy taxi-out delays in each of the post-Rule summers, relative to the pre-Rule summer of 2009.[32] Baseline and TDR-attributable changes in taxi-out delays of more than 60 minutes can be estimated using the same approach we used to calculate TDR-attributable cancellations. Table 12 provides the results of these calculations.

Table 12: Reductions in Summer Season >60-Minute Taxi-Out Delays vs. 2009 Baseline

Delay Day Group	2010	2011	2012
High	3,866	3,684	5,506
Low	411	1,206	1,189
TDR Attributable Reduction in Delays	**3,455**	**2,478**	**4,317**

*Excludes August 27–29, 2011 (Hurricane Irene)
Source: Econometrica tabulations of BTS on-time performance data.

Table 12 shows that the majority of the reductions in lengthy taxi-out delays occurred on High Delay Days. The reduced tarmac waiting time for 3,500 flights with taxi-out delays of more than an hour in summer 2010, 2,500 flights in summer 2011, and 4,300 flights in summer 2012 can be reasonably attributed to the TDR.

Analysis of the Effect of the TDR on Passenger Welfare

The GAO recommended that the DOT "fully assess the impact of the tarmac delay rule, including the relationship between the rule and any increase in cancellations and how they affect passengers and, if warranted, refine the rule's requirements and implementation to maximize passenger welfare and system efficiency." The GAO report did not quantify the impact, either favorable or unfavorable, on passenger welfare resulting from the increases in flight cancellations attributable to the TDR.

Passenger welfare is complicated and difficult to measure. The impact of a cancellation can be expected to vary greatly among passengers on a specific flight, depending on each individual's or travel party's location, schedule, and preferences, as well as the nature of the plans that have been disrupted. The impact of a particular flight cancellation on passenger welfare also depends on a set of factors that will vary from flight to flight, departure time to departure time, and day to day:

- Proactive cancellations announced well in advance of the departure time provide passengers with increased flexibility to make alternative arrangements and allow them to avoid the time losses and direct expenses associated with making the trip to the airport.

[32] It could be argued that the TDR has resulted in improvements in airline, airport, and air traffic control operations that are not limited to days when compliance with the TDR may be a salient issue. However, there are other plausible explanations that account for at least a share of the reduction in lengthy taxi-out delays, including improvements in airport ground operations and reduced congestion.

- Passengers on flights that are cancelled during peak travel periods or flights in large waves of cancellations in response to severe weather events will have to wait longer on average to rebook a flight.
- Cancellations that strand travelers in locations other than their originating or final destination (i.e., cancellations of connecting flights) reduce a passenger's ability to mitigate the adverse impact of the increased waiting time associated with rebooking the flight.
- Passengers on flights that taxi out, return to the gate, and are subsequently cancelled incur an additional loss of time, a portion of which is spent waiting in less comfortable conditions on an airport runway, rather than in the airport terminal.[33]

Notwithstanding these qualifications, we present some ways of considering the welfare issues. One core metric of impact on passenger welfare is the average increase in waiting time before a traveler disrupted by a flight cancellation can travel on a rebooked flight. Our review of the research literature indicates that the 2004–2005 estimate cited in the HDR regulatory evaluation (517 to 547 minutes, or 8.5 to 9 hours) underestimates the average increase in waiting time experienced by passengers disrupted by TDR-attributable cancellations, because increased load factors have reduced the spare capacity available to accommodate rebooked passengers. Moreover, load factors are typically highest during the summer months when most TDR-attributable cancellations occur.

On the other hand, the M&J estimate of a 21-hour average wait for disrupted air travelers to rebook flights overstates the true impact on passenger welfare. Their "passenger displacement model" assumes that 95 percent of travelers rebook flights on the same carrier from the same airport and assigns these passengers to the next available flights in this artificially limited queue.[34] Imposing this restriction increases the projected number of passengers who are unable to rebook a flight for the next day and must wait more than 24 hours to rebook, with an attendant increase in the average waiting time for all disrupted passengers.

Given the likelihood that these estimates represent lower bound and upper bound approximations of the average rebooking lag, using an intermediate value—somewhere in the range between 12 and 15 hours—may more adequately characterize the current range of traveler experiences after cancellations.

In addition to the quantity of any TDR-attributable changes in travel and waiting time, the conditions under which these increases or decreases in travel or waiting time are spent will also affect the per-hour value that should be used to monetize the benefits of decreases in delays and costs of increases in cancellations attributable to the TDR. For this reason, the hourly value of travel time saved (VTTS) provided in current DOT guidance ($42.10/hour for air travel) cannot be simply applied to the estimated quantity of increased or decreased waiting time, because

[33] Collectively, these factors explain the analytical focus on the TDR's impact on cancellations during the summer months, where aircraft load factors are higher and the onset of severe weather events is more difficult to predict.

[34] Joshua Marks and Darryl Jenkins, "Modeling Passenger Reaccommodation Time for Flight Cancellations in Airline Networks," GWU International Institute for Tourism Studies Aviation Program, June 30, 2010.

waiting may take place in relatively more or less comfortable conditions.[35] In particular, passengers on cancelled or delayed flights may be better able to perform other activities or pass the time more enjoyably in the terminal (rather than on a boarded plane) or by going to a home or a hotel (rather than remaining in the terminal).[36]

Public transportation research cited in the HDR regulatory evaluation reports estimates of the fraction of an hour of travel time that riders are willing to pay or accept to receive a higher or lower "level of service" (LOS).[37] The hourly premiums required to compensate passengers for riding in less comfortable conditions were used as proxies for the value of the welfare loss to air travelers of waiting on an aircraft experiencing a lengthy tarmac delay, rather than in an airport terminal. In addition to the uncertain comparability of the disruptions caused by flight cancellations and lengthy tarmac delays to conditions and delays on public transportation, the per-hour estimates cited do not vary with the time required to rebook a flight or the length of the taxi-out delay experienced.

The influence of the length of a travel delay on per-hour passenger welfare losses is directly addressed in recently updated European air travel delay valuation guidance developed by the University of Westminster.[38] The estimates presented in Table 13 show that there are sharp increases in the per-minute costs to passengers as the length of the delay increases, a result that is especially applicable to delays experienced by passengers who are already enplaned and experience a lengthy taxi-out delay.

Table 13: Cost to Passengers of Each Additional Minute of Travel Time by Delay Interval

Delay (minutes)	30	60	90	120	180	240	300
Low scenario	0.20	0.35	0.48	0.61	0.84	1.05	1.25
Base scenario	**0.34**	**0.58**	**0.79**	**0.99**	**1.36**	**1.70**	**2.04**
High scenario	0.41	0.71	0.96	1.20	1.66	2.08	2.48

Source: University of Westminster report for Eurocontrol Performance Review, March 31, 2011. Dollar amounts calculated at exchange rate of $1.42 per Euro, as of March 31, 2011.

Conversely, the per-hour welfare losses to passengers who have flights cancelled may actually decline with a longer rebooking time. Passengers who are rebooked on flights that depart within a few hours are likely to remain in the terminal, unless they become aware of a proactive cancellation before leaving for the airport. However, those who are rebooked on flights the next or following day (the longest delays) will leave the airport and are very likely to return home or spend the layover in more comfortable conditions with improved opportunities to engage in other business or personal activities of value.

[35] Office of the Secretary of Transportation, "Revised Departmental Guidance for the Valuation of Travel Time in Economic Analysis," September 28, 2011.
[36] In contrast, passengers who are onboard aircraft experiencing lengthy taxi-out delays have very limited options to alter the conditions under which they spend the additional time before departure.
[37] Todd Litman, "Valuing Transit Service Quality Improvements," Victoria Transport Policy Institute, November 24, 2011.
[38] Department of Transport Studies, University of Westminster, "European Airline Delay Cost Reference Values," Final Report (Version 3.2), March 31, 2011.

Based on our review of the research available, it may not be possible to develop satisfactory estimates of the impact of the TDR on passenger welfare during the post-Rule summers.

Conclusion

Based on the analysis and results presented in this report, the TDR has had some adverse and some beneficial impacts during the period analyzed**.** The TDR appears to have had an adverse impact on flight cancellations in the summer months of 2011. The TDR also appears to have had a smaller adverse impact on flight cancellations in 2010 and 2012 using one of the two methods developed for this analysis. However, the impact of the TDR on cancellations during the summers of 2010 and 2012 appears to have been minimal or non-existent using a second analytical approach presented in this report. In addition, the TDR appears to have reduced taxi-out waiting times for several thousand flights in each of the three post-Rule summers.

It is difficult to characterize the overall impact of TDR-attributable changes in cancellations and lengthy taxi-out delays on passenger welfare.

Appendix A: Comparison of TDR Study Impact Estimates

Table A-1: Tarmac Delay Rule (TDR) Study Methodology, Data, and Estimated Impact on Cancellations

Date	Study	Excess Cancellation Estimate	Post TDR Data Used	TDR-Attributable Cancellations*		
				2010	2011	2012
2009 (Dec.)	HDR regulatory evaluation to accompany Final Rule	2.8% of the flights in 2007-2008 that incurred a 3+ hour taxi-out delay	2007 - 2008	41	X	X
2010 (July)	Marks/Jenkins - report 1	20% of all flights cancelled on airport days with severe weather events plus equal number of follow-on cancellations	May 2010	5,200	X	X
2010 (Sept.)	DOT/Econometrica review of Marks/Jenkins report	Change in cancellations of flights w/ 2+ hour taxi-out delays from 2009	May - July 2010	none	X	X
2010 (Nov.)	Marks/Jenkins - report 2	Entire increase in cancellations from 2009 to 2010	May - Sept 2010	5,068	X	X
2011 (March)	DOT internal analysis	Change in number of cancellations on days w/ at least one >2 hour taxi-out delay	May - Oct 2010	none	X	X
2011 (March)	Marks/Jenkins - report 3	Change in % of all flights cancelled	May - Oct 2010	not est.	X	X
2011 (Sept.)	GAO Report	Increase in statistical probability of flight being cancelled at the gate or after taxi-out delay, estimated using logistic regression model	May - Sept 2010	not est.	X	X
2011 (Nov.)	Marks/Jenkins (now AAI) - report 4	Applied GAO % increases to all flight cancellations	May - Sept 2010, 2011	8,114	13,087	X
2014 (Jan.)	Econometrica - Approach 1	Increase in cancellations on days with > 1 taxi-out delay of > 2 hours, net of increase in cancellations on days with 0 or 1 of these delays	May - Sept 2010 - 2012	2,444	9,415	1,080
2014 (Jan.)	Econometrica - Approach 2	Increase in cancellations on 1/3 of days that had largest number of taxi-out delays of > 2 hours, net of increase in cancellations on 1/3 of days with lowest number of these delays	May - Sept 2010 - 2012	491	4,139	none

*Total flight cancellations during the May-September period were 33,532 in 2010, 46,989 in 2011, and 29,027 in 2012.

Source: Econometrica compilation of TDR impact study estimates, last revised August 7, 2013.

Appendix B: Excess Cancellations in June, July, and August

This appendix provides excess cancellation results for a sensitivity analysis performed for the post-Rule summer seasons that exclude the OTP data for May and September. In general, these estimates of TDR-attributable cancellations are consistent with those presented in the report.

Approach 1: TDR Impact Measured as the Number of Excess Cancellations on Days With More Than One Taxi-Out Delay of More Than 2 Hours

The first set of excess cancellation estimates were developed using cancellations on days with either zero or one taxi-out delay of more than 2 hours as the Low Delay Day baseline. All days with more than one lengthy taxi-out delay were included in the High Delay Day group.

Table B-1 shows the numbers of cancellations for these 2 groups in each summer season. [39]

Table B-1: June–August Departures and Cancellations by Delay Group (Approach 1)

Year	Delay Day Group	Scheduled Departures	Actual Cancellations	Adjusted Cancellations
2009	0/1	313,806	2,187	2,187
2009	>1	1,392,223	18,551	18,551
2010	0/1	641,006	5,810	2,844
2010	>1	1,050,686	16,252	21,535
2011*	0/1	520,757	5,456	3,288
2011	>1	1,047,836	19,993	26,564
2012	0/1	637,876	5,548	2,729
2012	>1	974,981	15,047	21,486

*Excludes August 27–29, 2011 (Hurricane Irene)
Source: Econometrica tabulations of BTS on-time performance data.

The adjusted cancellation estimates were then used to compute the number of excess cancellations on High Delay Days in each post-Rule summer shown in Table B-2.

[39] These calculations exclude the OTP data for the 3 days in August 2011 during which Hurricane Irene caused thousands of cancellations. Because there were no taxi-out delays of more than 2 hours, including these 3 days in the Low Delay Day group would unrealistically increase the number of baseline cancellations in 2011, resulting in large *negative* estimates of excess cancellations.

Table B-2: June–August Cancellations by Delay Day Group vs. 2009 Baseline (Approach 1)

Number of >2-Hour Taxi-Out Delays	2010	2011	2012
>1	657	1,101	542
0/1	2,984	8,013	2,935
Excess Cancellations Over Baseline	**2,327**	**6,912**	**2,393**

*Excludes August 27–29, 2011 (Hurricane Irene)

Based on this approach, 2,327 cancellations in June–August 2010, 6,912 cancellations in June–August 2011, and 2,393 cancellations in June–August 2012 can be reasonably attributed to the TDR. The 2010 estimate is nearly identical to that for the 5-month summer period analyzed in the report. The June–August 2011 and 2012 estimates are lower and higher, respectively, than the May–September estimates for the same summer seasons.

Approach 2: TDR Impact Measured as Excess Cancellations on High Delay Days Defined Independently for Each Summer Season

Table B-3 reports the numbers of more-than-2-hour taxi-out delays per day used to define the three delay day groups in each year.

Table B-3: Number of >2-Hour Delays Included in Each June–August Delay Day Group (Approach 2)

Delay Day Group	2009	2010	2011*	2012
Low	0-4	0	0	0
Moderate	5-29	1-8	1-17	1-6
High	>30	>8	>17	>6

*Excludes August 27–29, 2011 (Hurricane Irene)

Table B-4 shows the actual and adjusted numbers of cancellations for the Low, Moderate, and High Delay Day groups in each pre-Rule and post-Rule summer season.

Table B-4: June–August Departures and Cancellations by Delay Group (Approach 2)

Year	Delay Day Group	Scheduled Departures	Actual Cancellations	Adjusted Cancellations
2009	Low	592,096	4,434	4,434
2009	Moderate	549,699	6,415	6,415
2009	**High**	**564,234**	**9,889**	**9,889**
2010	Low	479,115	4,148	5,126
2010	Moderate	615,387	6,547	5,848
2010	**High**	**597,190**	**11,367**	**10,740**
2011*	Low	520,757	5,456	6,203
2011	Moderate	527,290	7,752	8,081
2011	**High**	**520,546**	**12,241**	**13,268**
2012	Low	532,814	4,640	5,156
2012	Moderate	548,739	5,718	5,728
2012	**High**	**531,304**	**10,237**	**10,871**

Table B-5 presents the calculations of excess cancellations using this approach.

Table B-5: June–August Cancellations by Delay Day Group vs. 2009 Baseline (Approach 2)

Delay Day Group	2010	2011	2012
High	851	3,379	982
Low	692	1,769	722
Excess Cancellations Over Baseline	**159**	**1,610**	**260**

*Excludes August 27–29, 2011 (Hurricane Irene)

Using this approach, there were 159 excess cancellations in June–August 2010, 1,610 in June–August 2011, and 260 in June–August 2012 that can be attributed to the TDR. The June–August 2010 and 2011 estimates of TDR-attributable cancellations are significantly lower than the estimates presented in the report for the same years that include May and September.