

TI 2013-056/VIII
Tinbergen Institute Discussion Paper



The User Costs of Air Travel Delay Variability

Paul Koster

Eric Pels

Erik Verhoef

*Faculty of Economics and Business Administration, VU University Amsterdam, and
Tinbergen Institute.*

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at <http://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Duisenberg school of finance is a collaboration of the Dutch financial sector and universities, with the ambition to support innovative research and offer top quality academic education in core areas of finance.

DSF research papers can be downloaded at: <http://www.dsf.nl/>

Duisenberg school of finance
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 8579

The user costs of air travel delay variability

BY PAUL KOSTER*, ERIC PELS AND ERIK VERHOEF

DEPARTMENT OF SPATIAL ECONOMICS VU UNIVERSITY

TINBERGEN INSTITUTE

SUMMARY – We derive the expected user costs of US domestic air travel delay variability taking into account scheduling behavior of travelers. Travelers do not only consider mean arrival delays, but also face scheduling costs because they arrive too early or too late at their destination. The model allows travelers to anticipate arrival delay variability by choosing an earlier flight. We show that the expected user costs of US air traffic delays are underestimated by 16% if arrival delay variability is ignored.

Keywords – air traffic delay, travel time variability, scheduling, value of reliability.

* Corresponding author. Department of Spatial Economics, VU University Amsterdam, De Boelelaan 1105 1081 HV Amsterdam, e-mail: p.r.koster@vu.nl. We like to thank participants of the Kuhmo-Nectar Conference 2011 and Eric Kroes and Hans Koster for comments and suggestions. Furthermore, we like to thank Martijn Smit of the Department of Spatial Economics of VU University Amsterdam for help with the preparation of the dataset. Any shortcomings or errors in the paper are of course our responsibility. Erik Verhoef and Paul Koster gratefully acknowledge the financial support from the ERC (Advanced Grant OPTION #246969).

1.0 Introduction

Unreliability of public and private transport systems can lead to high social costs. The valuation of reliability is therefore one of the four key topics included in the Second Strategic Highway Research Program (SHRP2) of the US government.¹ Also in Europe reliability of transport systems has gained increasingly attention after preliminary research has suggested that user costs of travel time variability may be substantial (OECD/ITF, 2010). This paper is an empirical contribution to the research on the valuation of reliable transport systems. It studies the impact of US air arrival delay variability on the expected user costs of air travelers. Up to now, little attention has been paid to the large scale empirical analysis of air travel delay variability. Earlier studies have considered travel delay variability for travelers going to the airport (Koster et al. 2011, Tam et al. 2008), but as far as we know no large scale empirical studies on arrival delay variability at the destination are available in the literature.

Travelers are not only concerned with mean arrival delay costs and prefer to arrive at their destination airport on time. Air travel delays potentially result in a costly disruption of travelers' activity schedules. Previous studies on US air travel delays included the user costs of mean arrival delay but neglected the costs of arrival delay variability. Our study is a first attempt to quantify these user costs.

Because airlines determine the scheduled arrival times of the flights, travelers need to account for the timetable when choosing a flight. Therefore, they face timetable constraints and cannot freely choose their arrival time, leading to scheduling costs.² Scheduling costs are considered as important and many studies have found significant effects for increases in flight frequency. For example, Douglas and Miller (1974) assume that the cost of arriving are symmetric around the preferred arrival time, and study the relationship between schedule delay costs and ticket prices. Anderson and Kraus (1981) analyze the relationship between schedule delay and demand empirically. They show that schedule delay has a

¹ The details of the program can be found on the website of the Transportation Research Part Board (www.trb.org).

² See for example Bates et al. (2001) for a discussion and Noland and Polak (2002)

significant impact on travel demand. Also Lijesen (2006) and Hess et al. (2007) find significant effects of schedule delay on travel demand using stated preference data.

In order to capture the dynamic behavioural response to arrival time variability, an explicit treatment of trip timing needs to be included in the model. This is done using the concept of schedule delay which was introduced by Vickrey (1969) and Small (1982) and later extended by Noland and Small (1995) to account for travel time variability. We refer to these scheduling preferences as α - β - γ scheduling preferences, where α is the value of travel time, β the value of arriving earlier than the preferred arrival time, and γ is the value of arriving later than the preferred arrival time.³

The main goal of this paper is to quantify the user costs of air travel delay variability. We use data on US domestic air traffic delays for the top 500 busiest origin-destination pairs. This dataset is unique since it provides very detailed information on scheduled arrival times and realized arrival times of US airlines. We find that, for reasonable parameter values, the costs of air traffic delays are underestimated by 16% if arrival delay variability is ignored. Section 2 introduces the stylized scheduling model that is suited for empirical application. Section 3 discusses the data, section 4 the results, and section 5 concludes.

2.0 The scheduling model

We extend earlier models of Fosgerau (2009) and Tseng et al. (2012) to make them more suitable for empirical analysis by taking into account that arrival times of flights are not equally spaced over the time of the day. This is important when arrival times are variable, since anticipating behaviour to earlier connections depends on how connections are spaced over the time of the day. If a connection is nearby the preferred arrival time, the probability

³Fosgerau and Engelson (2011) and Engelson and Fosgerau (2011) use other scheduling preferences proposed by Vickrey (1973) and Tseng and Verhoef (2008). They develop reduced form cost functions that take into account the restricted departure time decision when traveling with a scheduled service. They show that these scheduling preferences may lead to expected user costs that are linear in the variance of the delay distribution. This is a useful theoretical result because the variance is additive over links. Engelson and Fosgerau (2011) generalize this model. Which scheduling preferences are most appropriate for air travelers is an empirical question that will not be addressed in this paper.

that a traveler will choose this connection will be higher, meaning that there is an interaction between frequency and the expected costs of arrival time variability.

Consider the following set-up. For a given OD-pair, on a given day $v = 1 \dots V$, travelers choose between $N_v - 1$ connections labelled by $n_v = 1 \dots N_v$. It is assumed that $N_v \geq 2$, in order to have a possible trade-off between connections on day v . The headway H_{vz} is the difference in the scheduled arrival times of connections n_v and $n_v + 1$. On day v , there are N_v connections and therefore there are $N_v - 1$ values for the headway. We therefore define the variable $z = 1 \dots N_v - 1$, as the index for the headway. This setup therefore accounts for unequally spaced flights over the time of the day, since $H_{v1} \dots H_{v(N_v-1)}$ are not necessarily equal.

The scheduled arrival time is defined by sat_{n_v} , and the scheduled arrival time of the first connection (sat_1) is normalized to 0. It is assumed that all travelers have an exogenous preferred arrival time (pat) at their destination airport, which is somewhere in between the scheduled arrival times of connections 1 and N_v , meaning that $0 \leq pat \leq sat_{N_v}$. One might argue that from a system equilibrium perspective, locating the arrival time at the beginning or endpoint of the market is not optimal from an airline profit perspective. For simplicity this is ignored because analyzing a full equilibrium model is highly complicated in the current setting because one needs a multi-firm, multi-product formulation of the Hotelling model. Furthermore, one needs to account for the connectivity of the flights in the network. The pat is distributed over the time of day with a probability density function $\Phi'_v[pat]$.

Furthermore we assume that travelers have what we may call α - β - γ preferences (Vickrey, 1969; Small, 1982).⁴ The shadow costs of mean arrival delay are given by α , the shadow costs of arriving earlier than the preferred arrival time are given by β , and the shadow costs of arriving later than the preferred arrival time are given by γ . Empirical research for other travel modes usually finds that arriving late is more costly than arriving

⁴ Other scheduling preferences can be used (Vickrey, 1973; Tseng and Verhoef, 2008; Fosgerau and Engelson, 2011), but no empirical evidence for such preferences is available in the context of aviation. The model developed in this paper can easily be extended by using other scheduling preferences in equation (1). Tseng et al. (2012) and Börjesson and Eliasson (2011) also use these scheduling preferences.

early, meaning that $\gamma > \beta$, although we may add that the evidence for air travelers is mixed (Warburg et al., 2006; Lijesen, 2006; Hess et al. 2007).

The arrival delay D is distributed with a cumulative distribution function $F[D]$, and corresponding probability density function $F'[D]$. D can be negative, meaning that passengers may arrive earlier at their destination than the scheduled arrival time. Here it is assumed that the delays D come from the same underlying distribution, meaning that we do not account for seasonal changes in the arrival delay distribution. We standardize the delay to make the mean delay (μ) and the standard deviation of delay (σ) explicit. To that end, define $x=(D-\mu)/\sigma$, as the standardized delay, where x is distributed with a standardized distribution $G[x]$ and corresponding probability density function $G'[x]$. Throughout our analysis it is assumed that the distribution of arrival delay does not vary over the day, meaning that μ and σ are independent of the time of the day. This will lead to an underestimate of the expected user costs for certain times of the day, where μ and σ are larger and an overestimate for times that μ and σ are lower. However, these effects to some extent average out because travelers are by assumption distributed over the day.⁵ A traveler choosing connection n_v , having a delay D , and a scheduled arrival time sat_{n_v} , has a generalized travel cost $C[pat; D; n_v]$:⁶

$$(1) \quad C[pat; D; n_v] = p_{n_v} + \alpha \cdot D + \beta \cdot \text{Max}(0, pat - (sat_{n_v} + D))^\lambda + \gamma \cdot \text{Max}(0, sat_{n_v} + D - pat)^\lambda.$$

The first part of this equation p_{n_v} , is the cost for using the connection other than delay and scheduling costs. This includes the ticket price, the (monetized) costs of scheduled travel time, the service quality and frequent flyer miles. For the remainder of our analysis p_{n_v} is assumed to be constant over the day and therefore we normalize p_{n_v} to 0. Therefore, on a given day v , travelers base their choice of connection only on the delay and scheduling costs.

⁵ Note that we use 'to some extent' for two reasons. First, the mean and the standard deviation may have a nonlinear effect on the user costs and second: the mean and standard deviation could be related to the number of travelers.

⁶ Our analysis thus focuses on direct connections and does not take into account transfers.

The second part of equation (1) is the additional arrival time delay costs, which are the value of mean arrival delay savings multiplied by the delay. The third and the fourth part of equation (1) are the costs of not arriving at the *pat* multiplied with the corresponding shadow cost. The parameter λ captures possible non-linear effects in the schedule delay costs. For the main part of our analysis we assume $\lambda=1$, but in our sensitivity analysis we will consider a case where $\lambda > 1$, meaning that large delays are relatively more costly and small delays relatively less costly, compared to linear schedule delay cost. It may well be that a delay decreases the scheduling costs for some travelers that arrive early if D would be equal to 0, because they arrive closer to their *pat*. However, because α is in general found to be higher than β , individual trip cost will always increase if D increases. If scheduling costs are linear ($\lambda = 1$), and $\alpha < \beta$, travelers prefer a longer flight over landing too early.⁷

Following Noland and Small (1995), the expected user costs are then given by equation (2), where we substitute the standardized delay $D = \mu + \sigma \cdot x$ into equation (1) and integrate over all possible standardized delays:

$$(2) \quad EC[pat; \mu; \sigma; n_v] = \int C[pat; \mu + \sigma \cdot x; n_v] \cdot G'[x] dx.$$

We assume that travelers know the travel time distribution $F'[D]$. This is a strong assumption for air travelers since the knowledge of $F'[D]$ is mainly driven by experience and flight trips are made less regularly than car commuting trips. However, this analysis serves as a good benchmark because travelers that misperceive the probability distribution, will always have higher (or equal) expected user costs compared to the user equilibrium cost that will be derived from equation (2), since the choice of connection is not optimal anymore.⁸ This assumption secures that we make a conservative estimate of

⁷ For $\lambda > 1$, this may be the case for very early arrivals. With non-linear scheduling costs the units are important. For the mean delay and the schedule delay these are in hours.

⁸ Koster and Verhoef (2012) use a model for car travelers. However the argument also holds for travel with a scheduled service. In the absence of externalities, trip timing models based on expected utility theory, will give the lower bound of the user equilibrium expected cost. Any deviations because of (structural) misperceptions or probability weighting therefore result in non-optimal departure times and result in higher (or equal) expected user equilibrium costs. Other behavioural profiles might be considered, for example that travelers ignore delays completely in their trade-offs. The specification of the expected user costs function then will not change, but the solution for the optimal connection will change and expected user equilibrium costs will always be higher than in our analysis.

the costs of arrival delay variability. Travelers then choose the connection with the lowest expected cost from the set of N_v available connections on that day. This optimal connection is given by equation (3):⁹

$$(3) \quad n_v[pat, \mu, \sigma]^* = \operatorname{argmin}_{n_v \in \{1 \dots N_v\}} EC[pat; \mu; \sigma; n_v].$$

This optimal connection is a function of the pat , because travelers with a later pat are likely to prefer travelling with a later connection. More precisely, equation (3) is the solution to an integer optimization problem, and is a non-decreasing stepwise function in pat . Since we are not able to find an analytical solution without specifying the distribution of delays, we use numerical optimization to find $n_v[pat, \mu, \sigma]^*$, and use shorthand notation $n_v[.]^*$ to denote the optimal connection.¹⁰ The corresponding equilibrium expected cost for a given pat is then given by $EC[pat; \mu; \sigma; n_v[.]^*]$, where n_v in (2) is replaced by $n_v[.]^*$. The average (over individuals) total expected user equilibrium costs for day v are then given by the integral over the user equilibrium cost, multiplied by the probability density function of the preferred arrival times $\Phi'_v[pat]$, and a scale factor M_v that represents total demand over the day:

$$(4) \quad ATEC_v[\mu, \sigma] = M_v \cdot \int_0^{SAT_{N_v}} EC[pat; \mu; \sigma; n_v[.]^*] \cdot \Phi'_v[pat] dpat.$$

For simplicity we assume M_v is equal to 1. This means that (4) is the average expected user equilibrium cost over all preferred arrival times for day v .

To determine the costs of travel time variability, and to be able to compose it into various meaningful concepts, we define a number of benchmarks. The first benchmark we analyse concerns the cost level $ATEC_v[0,0]$, where it is assumed that $\mu=\sigma=0$ and there are no delays at all. This means that there are only deterministic scheduling costs. We make simplifying assumptions about the distribution of pat in order to keep the model tractable, and assume that the pat is uniformly distributed over the day, meaning that

⁹ The solution is unique except for the traveler who is indifferent between two connections. Assuming $\lambda=1$, the second derivative of (2) is given by: $\frac{\partial^2 EC[pat; \mu; \sigma; n_v]}{\partial n \partial n} = \frac{\beta + \gamma}{\sigma} \cdot G' \left[\frac{pat - n \cdot H - \mu}{\sigma} \right] \geq 0$.

¹⁰ For some specific distributions (binary/exponential) closed-form solutions are available (see for example Fosgerau and Karlström (2010) and Börjesson and Eliasson (2011)). But given the fact that delay distributions in our dataset do not have this particular shape, we estimate nonparametric distributions and use a numerical grid search to determine the optimal connection for each pat .

$\Phi'_v[pat]=1/sat_{N_v}$. In the sensitivity analysis we test several other distributions. A closed-form solution for the deterministic scheduling costs is derived in Appendix A for the non-linear scheduling cost function of (1), assuming $\mu = \sigma = 0$. For the linear case ($\lambda = 1$), the resulting average deterministic scheduling costs for day v is given by the following equation:¹¹

$$(5) \quad ATEC_v[0,0] = \sum_{z=1}^{N_v-1} \frac{1}{2} \cdot \frac{H_{vz}^2}{sat_{N_v}} \cdot \frac{\beta \cdot \gamma}{\beta + \gamma}.$$

These deterministic scheduling costs increase quadratically in the headways H_{vz} and increase if the schedule becomes more binding (increasing β and γ). If sat_{N_v} is fixed, for example due to governmental flight time restrictions, adding flights will decrease scheduling costs, as long as the scheduled arrival time of these new flights are not identical to that of existing flights (with available seats). Therefore, the value of increases in frequency also depends on the scheduled arrival times of other flights.¹²

As a second benchmark we analyse the cost under the assumption that there is no arrival time variability, meaning that we compare $ATEC_v[\mu, \sigma]$, with $ATEC_v[\mu, 0]$. In this case, travelers only consider the costs of an expected arrival time $sat_{nv} + \mu$, where μ is a deterministic constant. Some travelers then choose a different optimal connection, meaning that $n_v[pat, \mu, 0]^*$ is different from $n_v[pat, 0, 0]^*$. Average total cost without arrival time variability can be found by integrating over all possible preferred arrival times in a way similar to (4). Previous studies assumed a uniform distribution of pat , $\Phi'_v[pat]=1/sat_{N_v}$. In Appendix A we show that for a uniform distribution of preferred arrival times, $ATEC_v[\mu, 0]$ is given by:

$$(6) \quad ATEC_v[\mu, 0] = \alpha \cdot \mu + \sum_{z=1}^{N_v-1} \frac{1}{2} \cdot \frac{H_{vz}^2}{sat_{N_v}} \cdot \frac{\beta \cdot \gamma}{\beta + \gamma},$$

¹¹ This corresponds to the results of De Palma and Lindsey (2001), Fosgerau (2009) and Tseng et al. (2012). In their analysis, they assume $N_v=2$, $sat_1 = 0$ and $sat_2 = H$.

¹² Because (6) increases quadratically in the headways, it is more beneficial to add a flight where H_{vz} is larger, as long as preferred arrival times are distributed uniformly. For sat_{N_v} and the other scheduled flight arrival times fixed, the maximum possible reduction in scheduling costs is: $\frac{H_{max}^2}{4 \cdot sat_{N_v}} - \frac{1}{2} \cdot \frac{H_{max}^2}{sat_{N_v}} \cdot \frac{\beta \cdot \gamma}{\beta + \gamma}$, where H_{max} is the maximum headway on day v .

showing that with a uniform distribution of preferred arrival times the mean delay does not affect the average scheduling costs and therefore the average total expected cost increase linearly in the mean delay.

For every OD-pair we derive an aggregate measure of expected user costs including scheduling costs. To that end, we average (4)-(6) over all days of the year, where days with more flights are weighted heavier because it is likely that more travelers travelled on these days (we do not have passenger numbers per day). This average is given by $\overline{ATEC_{OD}[\mu, \sigma]} = \frac{1}{\sum_{v=1}^V N_v} \cdot \sum_{v=1}^V N_v \cdot ATEC[\mu, \sigma]_v$. We then define the average total costs of arrival time variability as $\overline{ATEC_{OD}[\mu, \sigma]} - \overline{ATEC_{OD}[\mu, 0]}$, and the average total costs of mean arrival delays as $\overline{ATEC_{OD}[\mu, \sigma]} - \overline{ATEC_{OD}[0, 0]}$.

3.0 Data

We use the “On-Time Performance” database of the year 2010, which includes scheduled and realized arrival times of domestic flights in the United States operated by airlines that carry at least 1% of the domestic flights in the US.¹³ The data is highly disaggregated and an individual observation corresponds to a single flight. Because there are systematic differences in the distributions of arrival delays between origins with the same destination, the analysis is performed separately for every OD-pair. The reason for these systematic differences in mean and standard deviation of arrival delays is that both at the origin and the destination delays occur. For example, adverse weather conditions or lack of airport capacity may occur both at the origin and the destination. Another important aspect of using OD-pairs is that this more closely matches the appropriate perspective for an air traveler. Travelers are not so much concerned with the aggregate arrival delay distribution at the destination airport, but rather face the delay distribution when travelling from their origin to their destination. We take the choice of origin and destination as given, meaning that we ignore the fact that in a multi-airport region travelers may choose between flights departing from multiple airports, or may change destination due to differences in arrival time delay costs. We include the top 500 OD-pairs, with the highest number of flights in

¹³ The dataset and the corresponding documentation can be downloaded at <http://www.transtats.bts.gov/>.

2010.¹⁴ These OD-pairs account for 37% of the total number of domestic flights in our dataset. The mean arrival delay (μ) and the standard deviation of arrival delay (σ) for the OD-pairs are given in Figure 1. The standard deviation of delays is increasing in the mean delay but the relation is less pronounced than for car travel (Peer et al. 2012). Some OD-pairs have negative mean delays while the mean arrival delay reaches a maximum at 0.26 hours. The standard deviation reaches values from 0.24 to 0.98 hours. There is a large dispersion of delays since $|\mu/\sigma|$ is smaller than one for all OD-pairs.

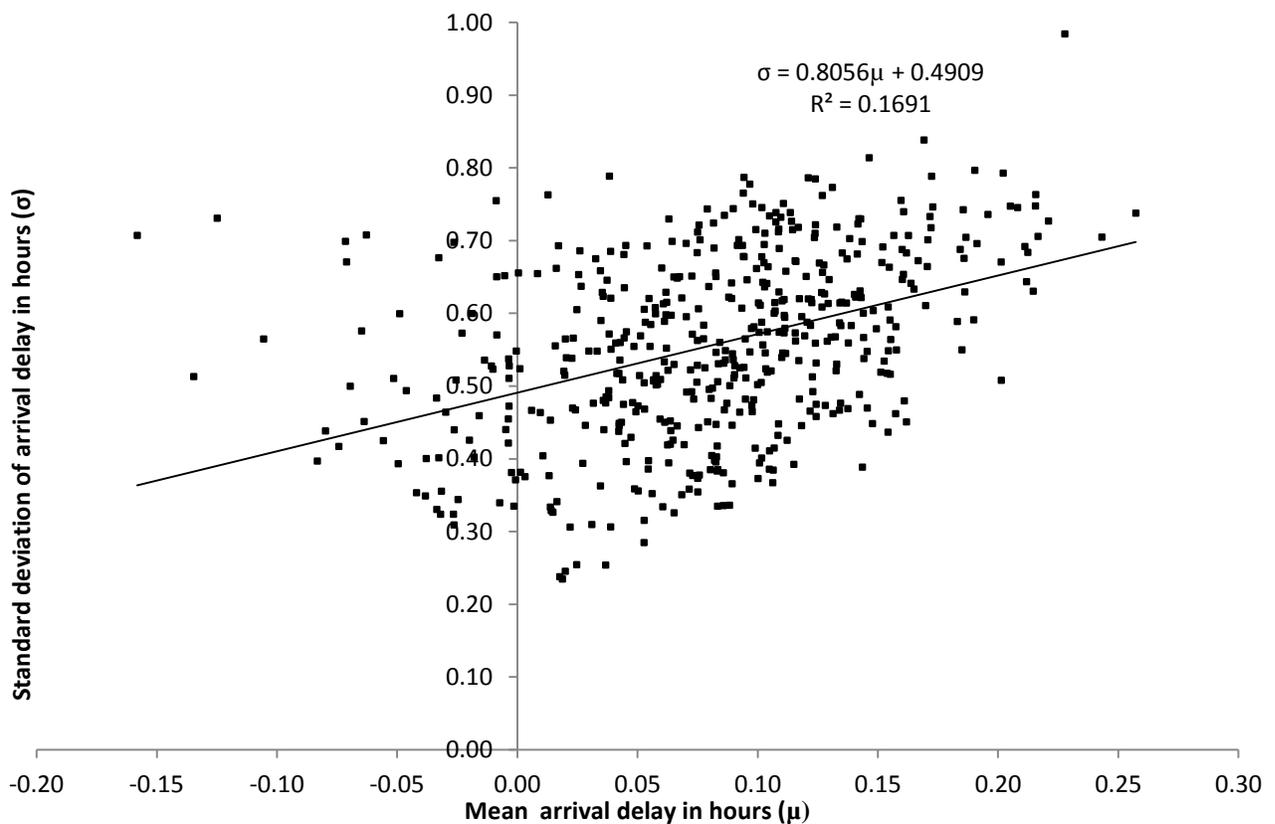


Figure 1 – Standard deviation of arrival delays (σ) as a function of mean arrival delay (μ).

Note: every dot in the figure corresponds to an OD-pair.

¹⁴ 3 OD pairs of the top 500 are omitted because these have a day with only one flight.

4.0 Results

4.1 Analysis for one OD-pair

We use a disaggregated level and perform the analysis for every OD-pair for every day v . In order to illustrate the steps of the model we show for one OD-pair the intermediate steps to obtain the results of equations (2)-(6). For this purpose we use the a flight from George Bush Intercontinental Airport (IAH) to Newark Liberty International Airport (EWR) as a representative OD-pair. This OD-pair is typical in terms of mean arrival delay (0.08 hours), and standard deviation of arrival delay (0.55 hours). Figure 2 shows the histogram of arrival delays. From the data the daily headways and scheduled arrival times are obtained and for each OD-pair a nonparametric density function of the delays is estimated using kernel smoothing.¹⁵ This avoids misspecification of the distribution which can result in an over or under estimation of expected user costs. It is assumed that the arrival delay distribution does not change over daytime and over the days of the year, so seasonal influences are ignored.

The next step is to calculate the total expected cost for one day. Therefore we need to assume values of schedule delay. We assume values of $\alpha = \$45$ per hour, $\beta = \$15$ per hour and $\gamma = \$30$ per hour. If flight arrivals are deterministic, and have a headway of 1 hour, our assumed values result in a value of headway of \$5 per/hour. As Lijesen (2006) shows, these values are in line with the existing literature on the valuation of frequency and schedule delay. Compared with estimations of Lijesen (2006, Table 5) our assumptions on the values of schedule delay can be regarded as conservative. In section 4.3, we show how the results depend on the assumed values of β and γ .

¹⁵ Estimation is done in Matlab using the `ksdensity` function using 500 equally spaced points. Bandwidths are optimal for a normal density (see Bowman and Azzilini, 1997). Results are available upon request.

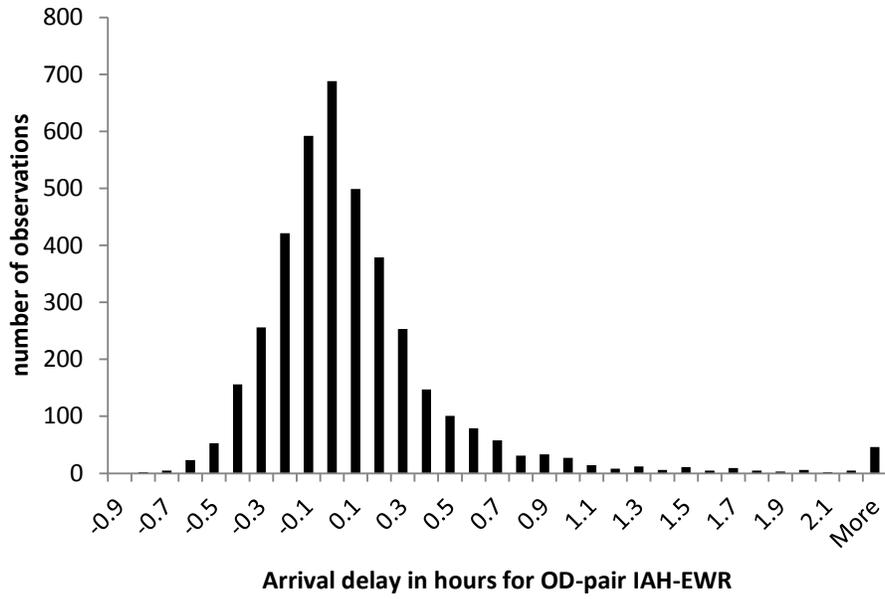


Figure 2 — Histogram of arrival delays(in hours) for OD-pair IAH-EWR.

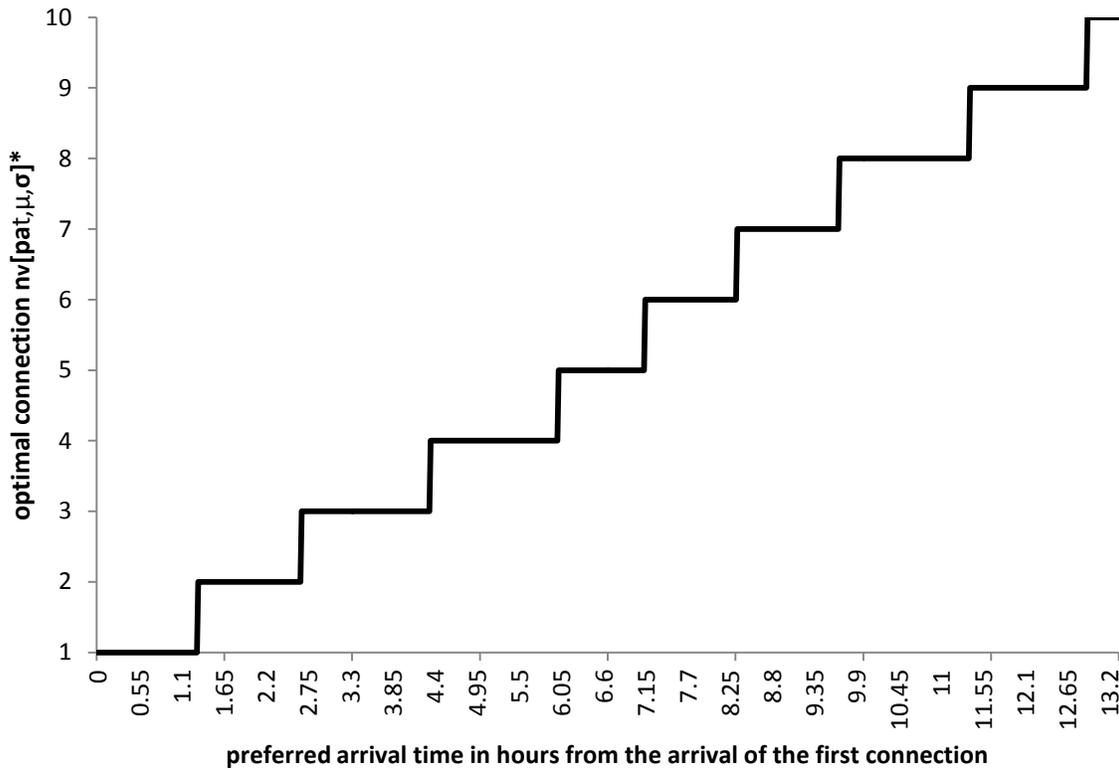


Figure 3 — $n_v[pat, \mu, \sigma]^*$ for a single day for OD-pair IAH-EWR.

Figure 3 shows the solution for the optimal connection as a function of the pat which was introduced in (3). On this day there are 10 flights scheduled, implying that $N_v=10$. The figure starts at the first connection and increases with 10 steps. The irregular pattern is caused by the fact that flights are not equally spaced over time-of-day, leading to unequal headways over the time of the day.

Figure 4 shows the expected user costs as a function of the preferred arrival time. The scheduled arrival times of the connections are indicated with vertical lines, where 0 indicates the scheduled arrival time of the first connection. The expected user cost functions for each connection as a function of the preferred arrival time are convex, and are given by the dashed lines.¹⁶ The traveler with $pat=0$ has the lowest expected cost for travelling with the first connection, and highest expected user costs for travelling with the last connection. The black line is the lower envelope of all the expected cost curves and gives the equilibrium expected cost $EC[pat; \mu; \sigma; n_v[.]^*]$ when travelers have chosen the optimal connection.

Figure 5 zooms in on this equilibrium cost in more detail. For this day, $ATEC_v[\mu, \sigma]$ equals \$14.65, $ATEC_v[\mu, 0]$ equals \$11.74 and $ATEC_v[0, 0]$ equals \$8.02. This means that the costs of arrival time variability are equal to \$2.92 (\$14.65-\$11.74), and the total costs of delays are equal to \$6.63 (\$14.65-\$8.02), which is 78% higher than what is found using the deterministic measure $\alpha \cdot \mu$ (\$3.72). Figure 5 also shows that the *minimal* user equilibrium costs are independent of H_{vz} . These costs are equal to the case where travelers could freely choose a departure time, and connections arrive continuously (see Fosgerau and Karlström, 2010). The optimal decision for this traveler then corresponds to the scheduled arrival time of the connection offered by the timetable.

The analysis per day enables us to show how the costs of arrival time variability change over the year. There may be fluctuations in scheduling costs, for example because airlines may schedule more flights during some days in the year and due to day of the week variations in the schedule. Figure 6 shows that there is some variation in $ATEC_v[\mu, \sigma]$ due to differences in the deterministic scheduling costs. That is, the difference between actual

¹⁶ Assuming $\lambda=1$, the second partial derivative of (2) is given by: $\frac{\partial^2 EC[pat; \mu; \sigma; n_v]}{\partial pat \partial pat} = \frac{\beta + \gamma}{\sigma} \cdot G' \left[\frac{pat - nH - \mu}{\sigma} \right] \geq 0$.

expected cost and cost that would be incurred if $\sigma=0$ is nearly constant, so that fluctuations over time must be due to fluctuations in deterministic cost components. The costs of arrival time variability $\overline{ATEC_{OD}[\mu, \sigma]} - \overline{ATEC_{OD}[\mu, 0]}$ have a mean of \$2.93, with a low standard deviation of \$0.04. A similar pattern arises for other OD pairs. Therefore it seems appropriate to report the results on a more aggregate level where we take the average over all the days of the year.

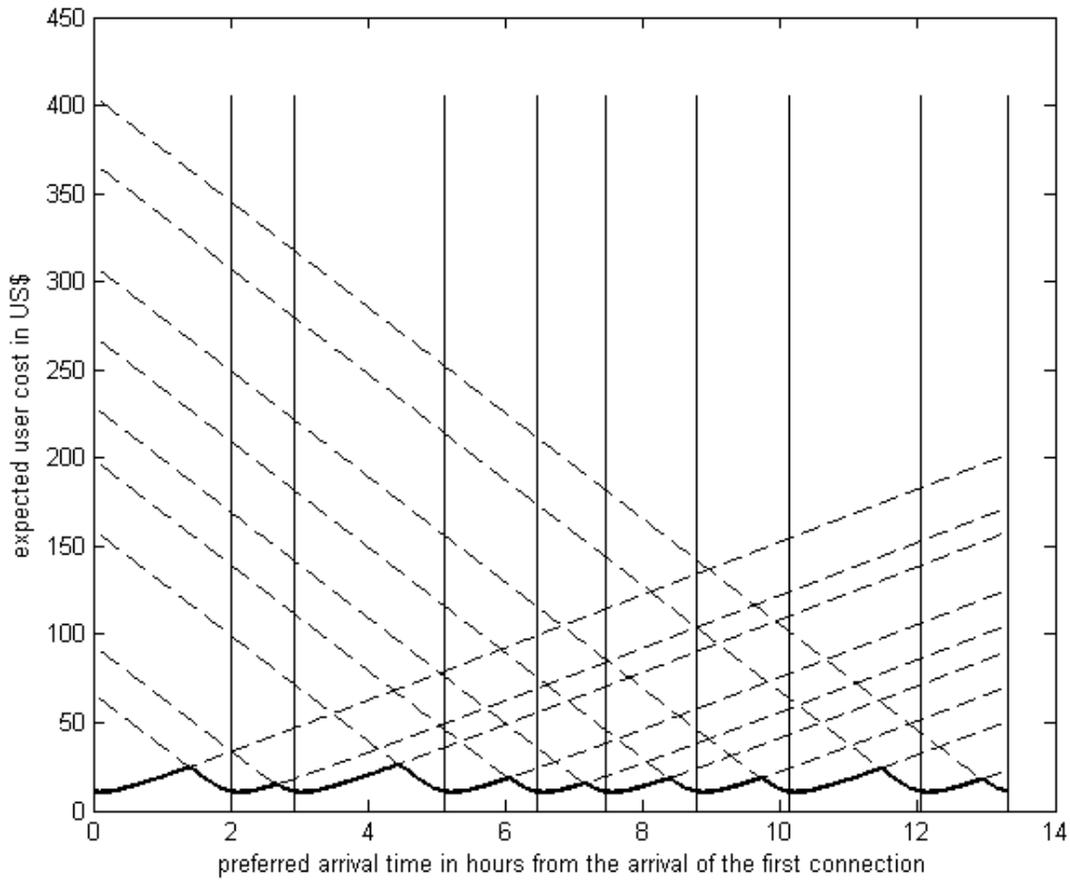


Figure 4 – Expected user costs functions (dashed lines) and equilibrium cost (black line) as a function of pat .

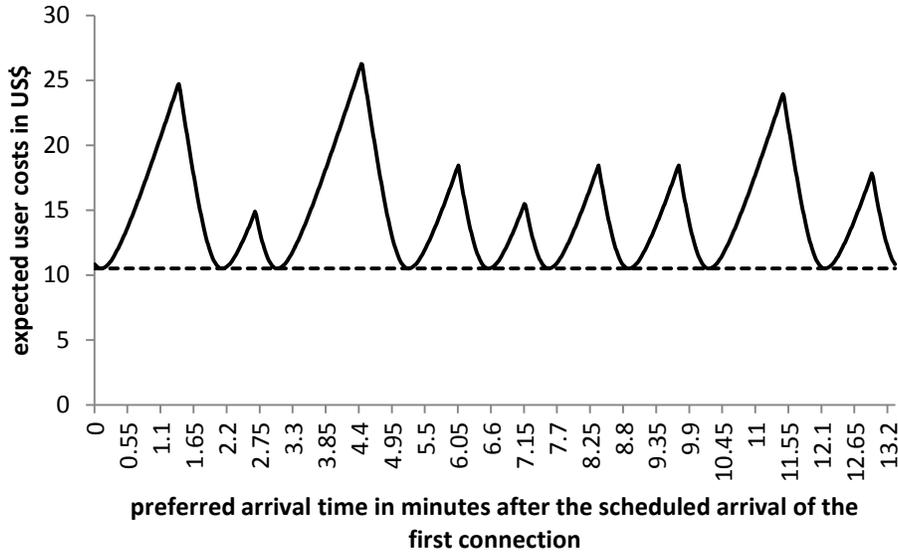


Figure 5 – User equilibrium cost (black line) and minimal equilibrium cost (dashed line) as a function of pat .

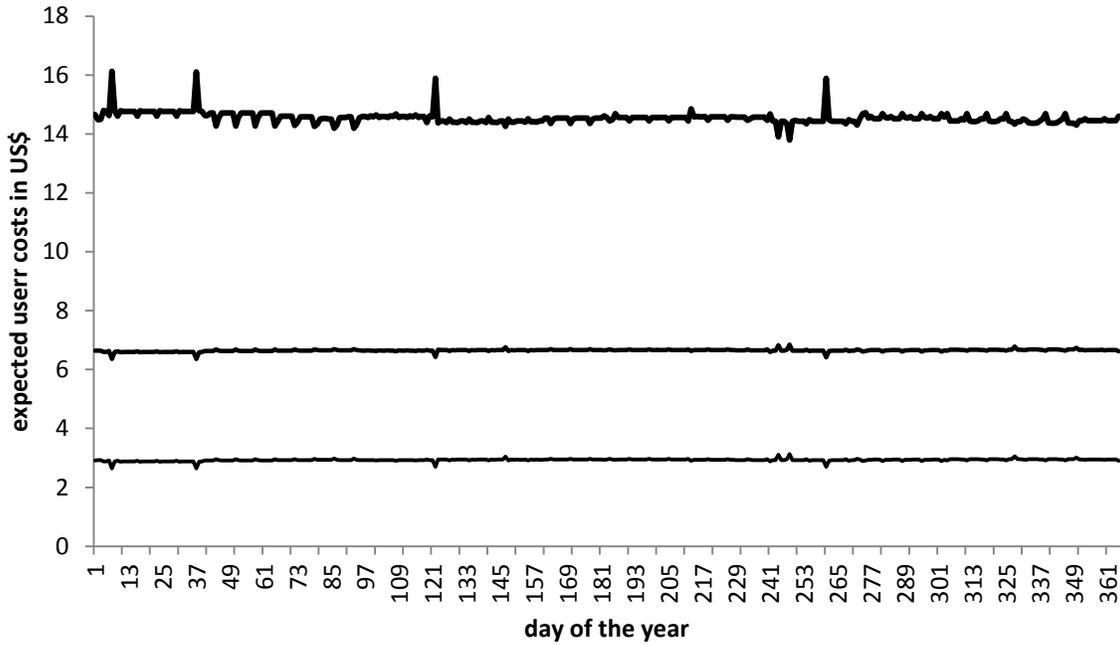


Figure 6 – Decomposition of total expected user costs. Upper line: $ATEC_v[\mu, \sigma]$, middle line: $ATEC_v[\mu, \sigma] - ATEC_v[\mu, 0]$, lower line: $ATEC_v[\mu, \sigma] - ATEC_v[0, 0]$ for every day of the year for the OD-pair IAH-EWR.

4.2 Analysis of the top 500 OD-pairs

We proceed with the full results of the top 500 OD-pairs. Therefore the analysis of section 4.1 is repeated, and we calculate the costs of arrival time variability for each OD-pair. Figure 7 shows the averaged values $\overline{ATEC_{OD}[\mu, \sigma]} - \overline{ATEC_{OD}[\mu, 0]}$ as a function of σ . The expected user costs of arrival time variability increase more than proportionally in the standard deviation of arrival delays and are in the range of 0.7-7.5 euro's per traveler.

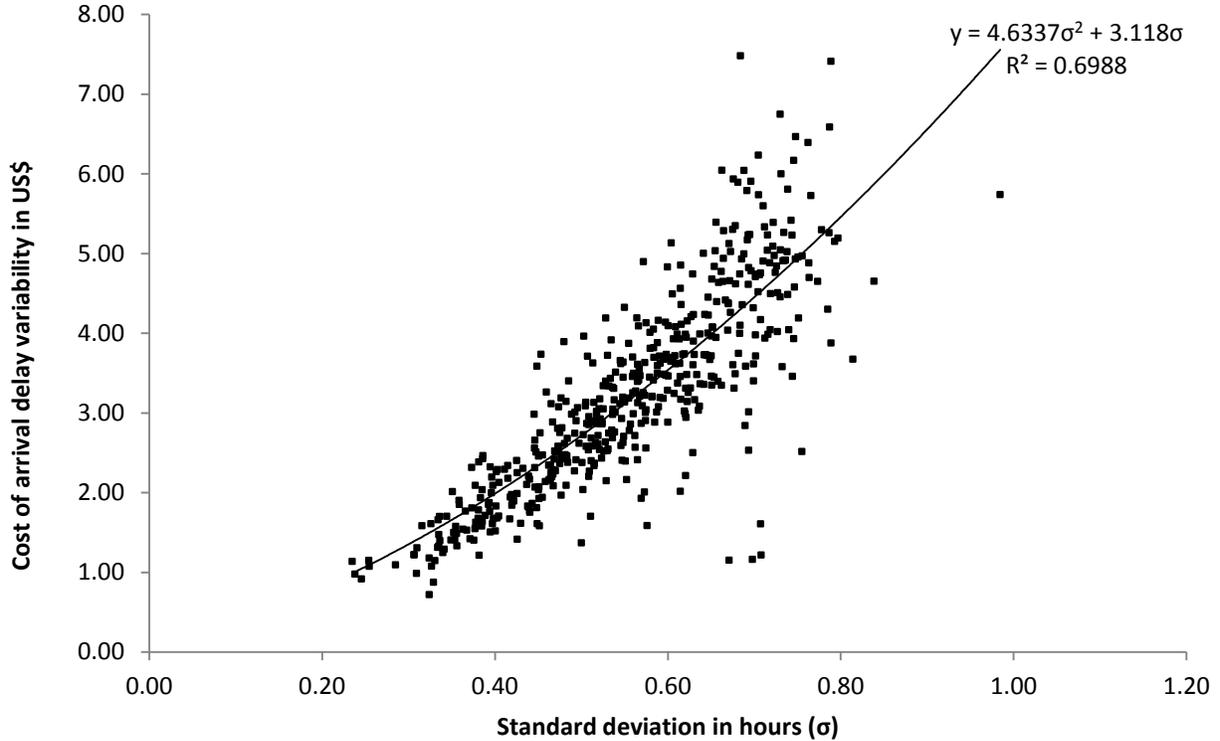


Figure 7 – Costs of arrival delay variability as a function of σ . Note: no constant is included because if $\sigma=0$, the costs of arrival time variability are 0.

Next, figure 8 shows the total expected user costs because of arrival delays as a function of the mean arrival delay μ . The first observation is that the expected user costs can be approximated well by a linear function of μ for a broad range of OD-pairs. This is good news for policy makers and airport congestion modellers, since the mean delay is then a good proxy for the total expected user costs because of delays and the complex dynamic choice problem of the travelers can be written in an empirically justifiable linear reduced form. Of course, this can only be used for the evaluation of policy measures that would not

disrupt this relationship. The lower linear line represents the equation $\alpha \cdot \mu$ and we thus see that the additional expected scheduling costs due to arrival time variability are substantial. The slope of the lower line is by assumption \$45. If we include expected scheduling costs, we find a slope of \$52, meaning that the expected user costs of arrival delays are underestimated by 16% if variable arrival times are ignored.

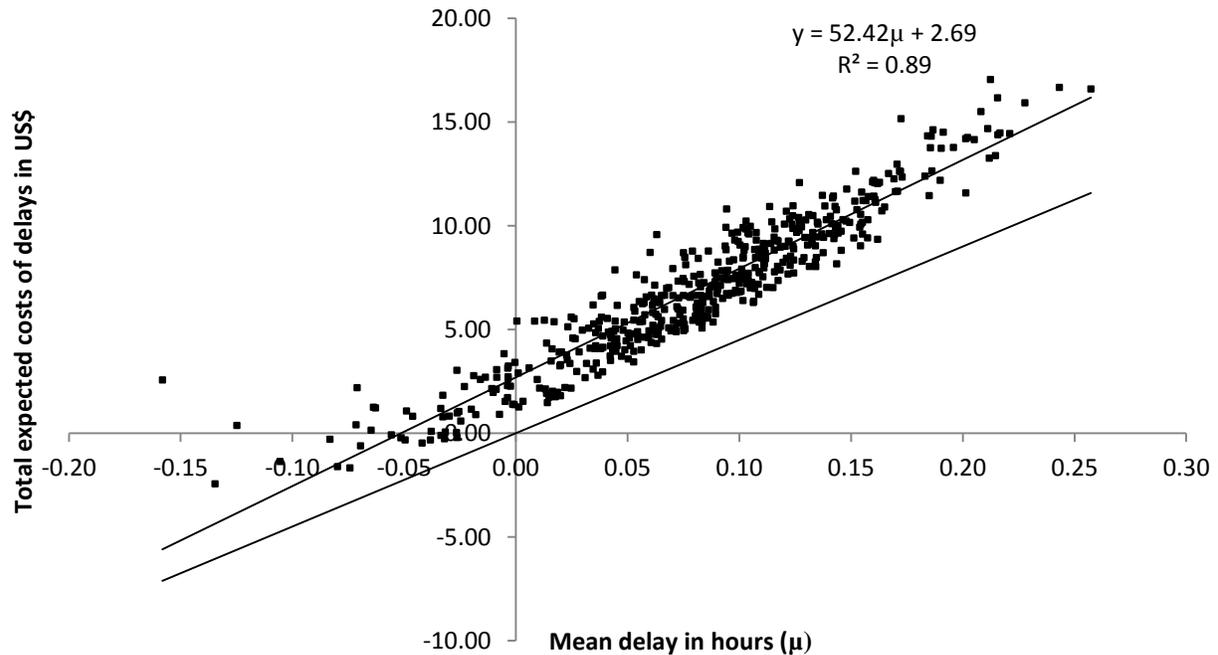


Figure 8— Average expected user costs because of arrival delays as a function of mean arrival delay. Note: the slope of the lower line is given by 45μ .

The linearity result is persistent for different distributions of preferred arrival times and values of λ .¹⁷ This means that for small changes of the mean delay the slope of the trend line can be interpreted as the ‘implied’ value of mean delay savings, meaning that it also includes the expected schedule delay cost. The lower line indicates the cost when only the mean delay is included in the user cost function. For small changes of the mean delay, the reduced form may be used as an approximation.

Therefore, the effects of air travel delays increase expected user costs more than previously thought, and as a consequence the potential welfare gains obtained from congestion pricing schemes as proposed by Brueckner (2002), Mayer and Sinai (2003) and calculated by

¹⁷ See appendix B and the discussion in section 4.3.

Morrison and Winston (2007) may be larger if arrival time variability is included in the user cost function, and variability would be decreasing if mean delays are reduced. Still, if the estimate of the value of travel delay savings (α) is based on revealed preference (RP) data, it will likely pick up the expected scheduling costs as well, meaning that the current estimates of congestion cost may already incorporate expected scheduling costs if α is based on RP analysis.

4.3 Discussion of the assumptions¹⁸

4.3.1 Assumption on equality of prices over time of the day

Throughout our analysis we assumed that the monetary cost p_{n_v} for other components than delays is constant over time-of-day for a given day v . If airlines do not differentiate their ticket prices over the time of the day, absolute fare levels do not have an effect on our estimate. This is probably not a realistic assumption since airlines may differentiate their prices over time-of-day in order to maximize their profits, or travelers may have a preference for a certain airline because of frequent flyer miles or other quality differences. If p_{n_v} is not constant over time of day, the choice of the optimal connection, as given by (3), will also depend on the other cost components $p_1 \dots p_{N_v}$. Given the assumption that p_{n_v} is constant over the time of the day, the decision to choose the optimal connection $n_v[.]^*$ is solely based on delay cost. Any other connection that is chosen will therefore by definition raise the expected user costs because of delays; otherwise $n_v[.]^*$ is not optimal. If p_{n_v} is not equal over time of the day, travelers choose $n_v[.]^*$ or a different connection. If they choose $n_v[.]^*$, the expected costs of delays is equal to what we found. If they choose another connection the expected costs of delays are higher. This means our estimate of the costs of delays is conservative since including the p_{n_v} in the cost function, would certainly increase equilibrium (expected) cost of delays, since the decision of the travelers is no longer optimal in terms of average and schedule delay cost.

¹⁸ Detailed numerical results of this section are available upon request.

4.3.2 Non-linear scheduling cost function

For non-linear schedule delay functions it is a-priori unknown if the equilibrium cost will increase or decrease for other values of λ . For given values of β and γ , if $\lambda > 1$, the scheduling costs will be lower than in the linear case for arrivals closer to pat . More precisely: if the arrival is less than an hour from pat , the scheduling costs will be lower. For arrivals further away (more than an hour from pat), the scheduling costs will be higher. Therefore, the effect of non-linear scheduling preferences depends on the timetable and the distribution of delays. For example, Li et al. (2012) study departure time choices of car drivers with non linear utility functions.

To show the numerical effect of non-linear scheduling preferences on the results, we re-estimated the user costs assuming a convex scheduling cost function, with a value of $\lambda = 1.3$ in equation (1) Compared to the model with linear scheduling costs, the costs of arrival time variability are on average 24% higher implying that assuming $\lambda = 1$ results in a conservative estimate of the average user costs. Even when the scheduling cost function is nonlinear, the expected user costs because of arrival delays can still well be approximated by a linear function of mean arrival delay. This is shown in figure 12 in Appendix B.

4.3.3 Distribution of preferred arrival times

Third, we test how the results depend on the assumed probability density function of the preferred arrival times. A priori there is no clear cut theoretical answer how a different distribution of pat affects our results. Some recent empirical evidence of Brey and Walker (2011) for domestic air travelers suggests that the distribution of preferred arrival times is not uniform and may follow a bimodal pattern over the time of the day. Therefore we perform some numerical sensitivity checks to see how the results are affected.

To keep the analysis tractable, we consider three other non-uniform distributions of pat . First, we assume that the probability is linearly *decreasing* over time-of-day, meaning that a larger share of travelers prefers to arrive in the morning. This may be typical for business travelers who usually prefer to arrive in the morning (Warburg et al., 2006). Second, it is assumed that the probability is *increasing* in time-of-day, meaning that a larger proportion

of travelers prefer to arrive in the evening. Third, it is assumed that the probability is symmetrically *U-shaped*, meaning that more travelers prefer to arrive in the morning and evening than near noon. The results of Figure 8 hardly change, and therefore the calculated average expected user costs due to delays are rather independent of the assumptions on the distribution of preferred arrival times. The corresponding figures can be found in Appendix B.

4.3.4 Values of schedule delays and proportional heterogeneity of preferences

The costs of arrival time variability depend on the assumed values of schedule delay. Of course one can run the same model again for different assumptions on α , β and γ . But if α , β and γ change in the same proportion, the results derived in this paper only change in an absolute sense, while the relative contribution of arrival time variability to total costs remains the same. This type of heterogeneity is elsewhere referred to as “proportional heterogeneity”, and may be caused by the fact that there is heterogeneity in the marginal utility of income, causing α , β and γ to vary in fixed proportions (van den Berg and Verhoef, 2011). This can be seen from the cost function of equation 1. Suppose that we have different groups in the population. If we multiply the assumed willingness to pay values α , β and γ for group i with a corresponding constant $k_i > 0$, the expected user costs are given by equation 7:

$$(7) \quad EC[pat; \mu; \sigma; n_v] = k_i \cdot \int C[pat; \mu + \sigma \cdot x; n_v] \cdot G'[x] dx.$$

This shows that the expected user costs are homogeneous of degree 1, and therefore the user equilibrium cost will shift with a fixed constant and the relative contribution of the costs of arrival time variability remains constant for all $k_i > 0$.

Suppose that we keep α constant and that we multiply β and γ by k_i . Because it is assumed that μ is equal over the time of the day, the choice of the optimal connection only depends on the ratio of γ/β . Therefore the solution of the choice of the optimal connection will not change. If the base values we chose for β and γ are multiplied by k_i , the costs of arrival time variability will be k_i times higher. Following the result of figure 8, this means

that the expected user costs for group i with corresponding k_i can be written in reduced form as: $(\alpha + k_i \cdot (52.42 - 45)) \cdot \mu$, where $k_i = 1$ corresponds to the result of figure 8.

5.0 Final comments and discussion

This paper showed that air travel delay variability for US domestic air travel, may raise the expected user costs of delays of air travelers with 16%, given our central assumptions on the value of delay savings and scheduling preferences. We view this as a conservative estimate, and we found for a broad range of origin-destination pairs that expected user costs because of air travel delays can be well approximated by a linear function of the mean delay. The fact that the mean delay is a very good proxy for the expected costs because of delays has been found by other research on car travel as well (Koster et al. 2011; Koster and Verhoef, 2012). Without repeating the arguments of the airport congestion pricing debate, our result strengthens the argument for introducing airport congestion pricing if delays entail externalities, since the welfare losses due to delays are higher than when only mean delay is taken into account in the user cost function, depending on whether the applied value of time already implicitly reflects the costs of arrival time variability in so far as correlated with expected travel time.

However it is not clear if current *revealed preference* studies into the value of time in aviation already implicitly include the expected costs of schedule delay in their estimation, because expected scheduling costs are so closely related to the mean delay cost. Our model might then explain the high values of times that are sometimes found in RP studies. If this is the case, current estimates of delay cost are more likely to be appropriate. Therefore there is a need for good estimates of the values of schedule delay and mean delay savings for air travelers, using revealed and stated preference data, in order to better disentangle the different cost components, which in turn may help to better prioritize policies that affect mean delays and variability in different ways.

Future studies may investigate in more detail how the mean delay and standard deviation of delays are related to congestion on the origin and the destination airport and the rest of the network, to gain more insights in the empirical relation between airport

congestion and arrival time variability. For example, Peer et al. (2012) investigate such relationships for car travelers.

We did not study the prediction of air travel delays in a larger equilibrium model. It would be interesting to study how ticket prices are related to headways and arrival time variability. In our analysis we assumed that the price of the connections is constant over time of day. As argued before, this is not a problem for the purpose of this paper, since it results in a lower bound estimate of the expected user costs because of arrival delays.

Appendix A Derivation of deterministic scheduling costs

A.1 Uniform *pat* distribution

We assume that the *pat* distribution is uniform and assume that $p_{nv} = 0$ and $\mu=0$. In the deterministic case, travelers then always choose between two connections. Connection z has a normalized scheduled arrival time 0, and connection $z+1$ has a scheduled arrival time H_{vz} . The *pats* are distributed between 0 and H_{vz} with probability $1/sat_{Nv}$. The cost functions are:

$$(A.1) \quad C_z = \beta \cdot pat^\lambda,$$

$$(A.2) \quad C_{z+1} = \gamma \cdot (H_{vz} - pat)^\lambda.$$

The next step is to determine for which values of *pat* travelers switch from connection z to $z + 1$. Solving for the switching *pat* gives:

$$(A.3) \quad pat^* = H_{vz} \cdot \frac{\gamma^{\frac{1}{\lambda}}}{\beta^{\frac{1}{\lambda}} + \gamma^{\frac{1}{\lambda}}}.$$

This means that all travelers on the interval $[0;pat^*]$ choose to travel with connection z . All travelers on the interval $[pat^*;H_{vz}]$ travel with connection $z + 1$. As long as $\gamma > \beta$, an increase in λ will lead to a decrease in *pat*^{*} meaning that more travelers use connection z .¹⁹ The average (over preferred arrival times) deterministic scheduling costs are then given by the integral over all preferred arrival time in between 0 and H_{vz} :

$$(A.4) \quad DSC_{vz} = \int_0^{pat^*} C_z \cdot \frac{1}{sat_{Nv}} dpat + \int_{pat^*}^{H_{vz}} C_{z+1} \cdot \frac{1}{sat_{Nv}} dpat.$$

$$(A.5) \quad DSC_{vz} = \frac{\gamma \cdot \left(\frac{H_{vz} \cdot \beta^{\frac{1}{\lambda}}}{\beta^{\frac{1}{\lambda}} + \gamma^{\frac{1}{\lambda}}} \right)^{1+\lambda}}{sat_{Nv} \cdot (1 + \lambda)} + \frac{\beta \cdot \left(\frac{H_{vz} \cdot \gamma^{\frac{1}{\lambda}}}{\beta^{\frac{1}{\lambda}} + \gamma^{\frac{1}{\lambda}}} \right)^{1+\lambda}}{sat_{Nv} \cdot (1 + \lambda)}.$$

De Palma and Lindsey (2001), Fosgerau (2009) and Tseng et al. (2012) assume $\lambda = 1$ and $sat_{Nv} = H$:

¹⁹ Note that $\frac{\partial pat^*}{\partial \lambda} = \frac{H_{nv} \cdot (\beta\gamma)^{\frac{1}{\lambda}} \cdot \text{Log}[\frac{\beta}{\gamma}]}{(\beta^{\frac{1}{\lambda}} + \gamma^{\frac{1}{\lambda}})^2 \cdot \lambda^2} < 0, \text{ if } \gamma > \beta.$

$$(A.6) \quad DSC_{vz} = \frac{\gamma \cdot \left(\frac{H_{vz} \cdot \beta}{\beta + \gamma}\right)^2}{2 \cdot sat_{Nv}} + \frac{\beta \cdot \left(\frac{H_{vz} \cdot \gamma}{\beta + \gamma}\right)^2}{2 \cdot sat_{Nv}} = \frac{H_{vz}^2}{2 \cdot sat_{Nv}} \cdot \frac{\beta \cdot \gamma}{\beta + \gamma}.$$

For a uniform distribution of preferred arrival times the mean delay does not affect the scheduling costs and the average costs of mean delay are then given by equation (A.7):

$$(A.7) \quad ATEC_v[\mu, 0] = \alpha \cdot \mu + \sum_{z=1}^{N_v-1} DSC_{vz}.$$

The main body of the analysis assumes that schedule delay costs are linear, so $\lambda = 1$. We illustrate the calculation of scheduling costs for other *pat* distributions for the case that $\mu = 0$ and $\sigma = 0$. The switching *pat* does not change, but the average costs of mean delay will change. If we assume a linearly decreasing probability for the preferred arrival times, $\Phi_v'(pat) = \frac{2 \cdot (sat_{Nv} - pat)}{sat_{Nv}^2}$. The deterministic scheduling costs are then given by equation

(A.8):

$$(A.8) \quad DSC_{vz} = \frac{H_{vz}^3 \cdot \beta \cdot \gamma \cdot (2 \cdot \beta + \gamma)}{3 \cdot sat_{Nv}^2 \cdot (\beta + \gamma)^2}.$$

If we assume a linearly increasing probability for the preferred arrival times, $\Phi_v'(pat) = \frac{2 \cdot pat}{sat_{Nv}^2}$. The deterministic scheduling costs are then given by the following equation:

$$(A.9) \quad DSC_{vz} = \frac{H_{vz}^3 \cdot \beta \cdot \gamma \cdot (\beta + 2 \cdot \gamma)}{3 \cdot sat_{Nv}^2 \cdot (\beta + \gamma)^2}.$$

If we assume a U-shaped probability for the preferred arrival times, with a minimum probability at $\frac{sat_{Nv}}{2}$, the distribution of preferred arrival times is given by: $\Phi_v'(pat) = \frac{12}{sat_{Nv}^3} \cdot \left[pat - \frac{sat_{Nv}}{2}\right]^2$. The deterministic scheduling costs are then more cumbersome and are given by the following equation:

$$(A.10) \quad DSC_{vz} = \frac{3 \cdot H_{vz}^2 \cdot \beta \cdot \gamma}{2 \cdot sat_{Nv} \cdot (\beta + \gamma)} - \frac{2 \cdot H_{vz}^3 \cdot \beta \cdot \gamma \cdot (\beta + 2 \cdot \gamma)}{sat_{Nv}^2 \cdot (\beta + \gamma)^2} + \frac{H_{vz}^4 \cdot \beta \cdot \gamma \cdot (\beta^2 + 3 \cdot \beta \cdot \gamma + 3 \cdot \gamma^2)}{sat_{Nv}^3 (\beta + \gamma)^3}.$$

The results of this appendix show that the marginal social costs of headway depends both on the curvature of the scheduling function (captured by λ) and the distribution of preferred arrival times.

Appendix B Results for different preferred arrival time distributions and non-linear schedule delay

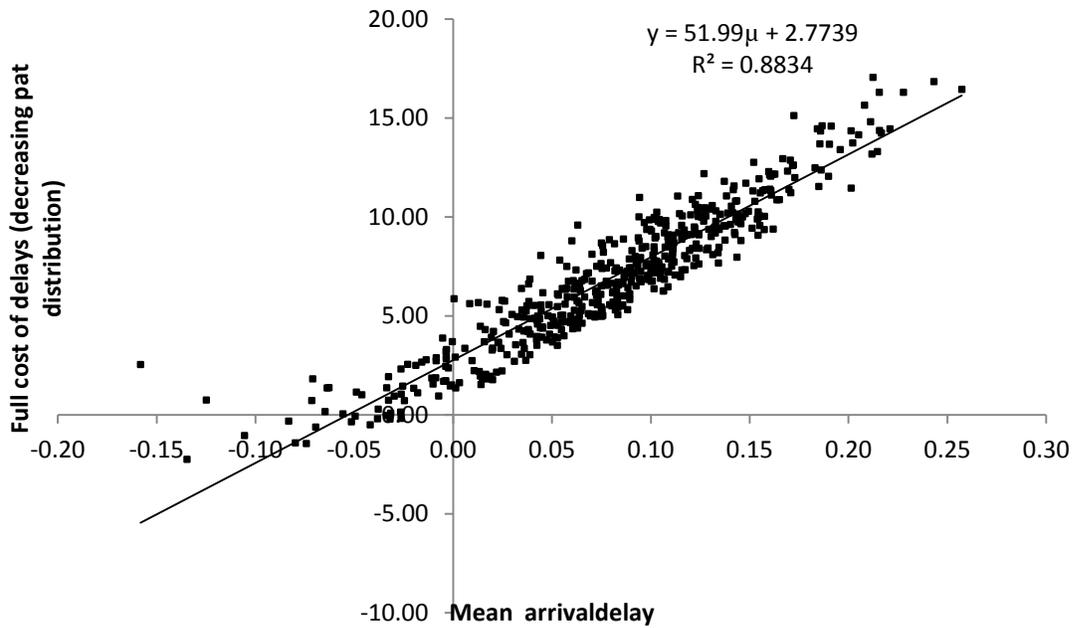


Figure 9— Average expected user costs because of arrival delays as a function of mean arrival delay for a decreasing pat distribution.

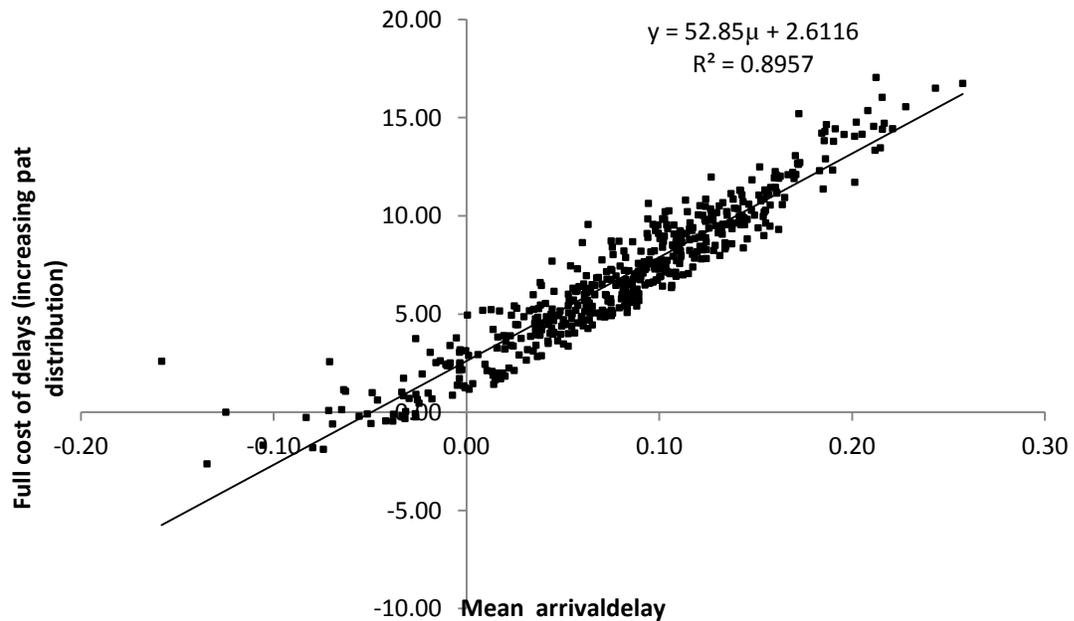


Figure 10— Average expected user costs because of arrival delays as a function of mean arrival delay for an increasing pat distribution.

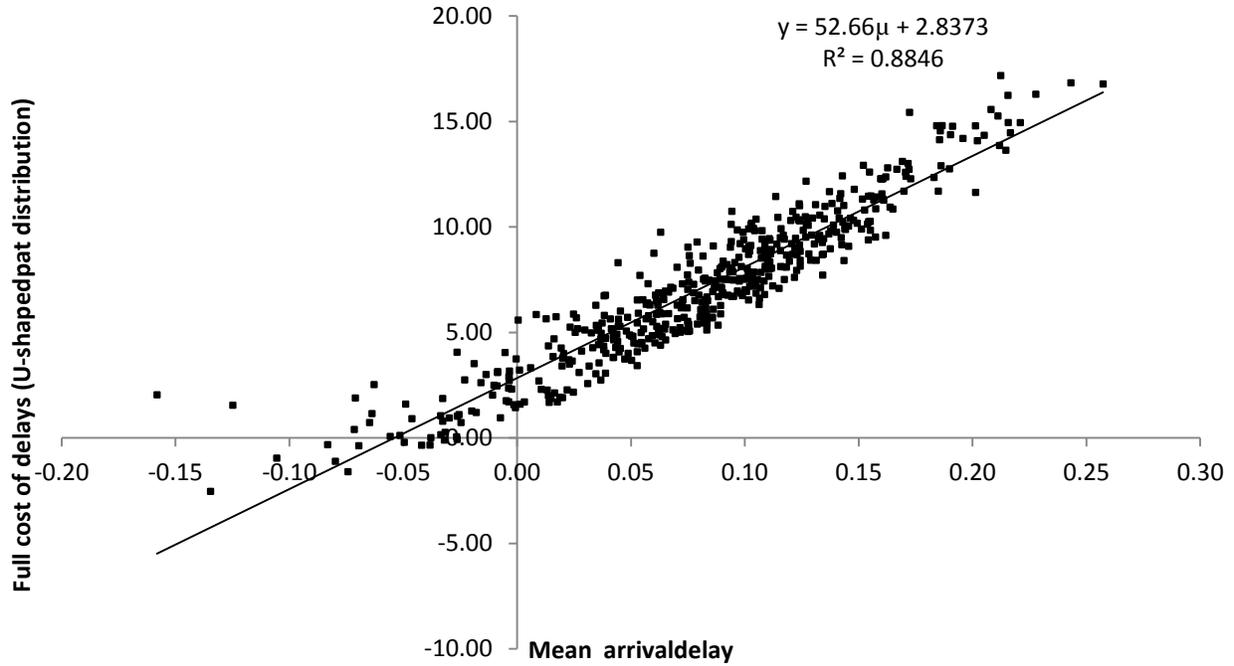


Figure 11— Average expected user costs because of arrival delays as a function of mean arrival delay for a U-shaped pat distribution.

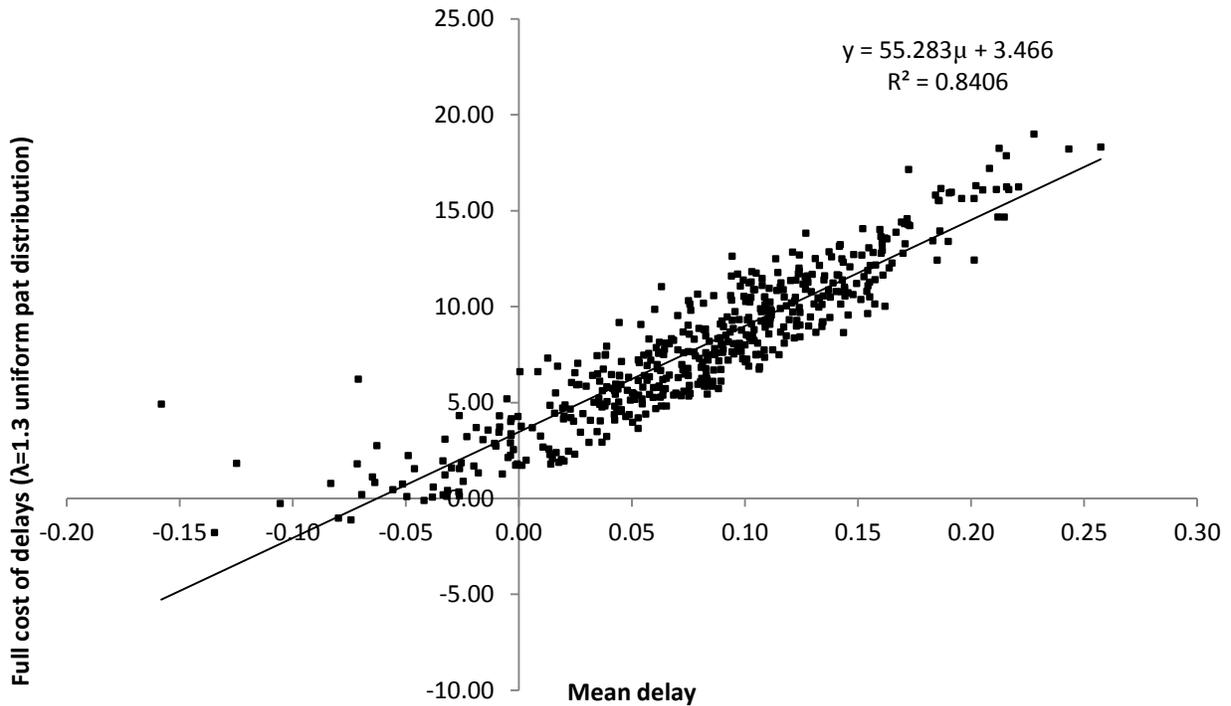


Figure 12— Average expected user costs because of arrival delays as a function of mean arrival delay for a $\lambda=1.3$ and a uniform distribution of preferred arrival times.

References

- Anderson, J.E., Kraus, M. 1981. Quality of Service and the Demand for Air Travel. *Review of Economics and Statistics* 63 (4), 533-540.
- Bates, J., Polak, J., Jones, P., Cook, A. 2001. The valuation of reliability for personal travel. *Transportation Research Part E* 37 (2-3), 191-229.
- Börjesson, M., Eliasson, J. 2011. On the use of average delay as a measure of train reliability. *Transportation Research Part A* 45 (3), 171-184.
- Brueckner, J.K. 2002. Airport Congestion When Carriers Have Market Power. *American Economic Review* 92 (5), 1357-1375.
- Carrion, C., Levinson, D. 2012. Value of travel time reliability: a review of current evidence. *Transportation Research Part A* 46 (4), 720-741.
- De Palma, A. and Lindsey, R. 2001. Optimal timetables for public transportation. *Transportation Research Part B* 35 (8), 789-813.
- Douglas, G.W. and Miller, J.C. 1974. Quality Competition, Industry Equilibrium and Efficiency in the Price-Constrained Airline Market. *American Economic Review* 64 (4), 657-669.
- Engelson, L., Fosgerau, M. 2011. Additive measures of travel time variability. *Transportation Research Part B* 45 (10), 1560-1571.
- Fosgerau, M. 2009. The marginal social costs of headway for a scheduled service, *Transportation Research Part B* 43 (8-9), 813-820.
- Fosgerau, M., Engelson, L. 2011. The value of travel time variance. *Transportation Research Part B*, 45 (1), 1-8.
- Fosgerau, M., Karlström, A. 2010. The value of reliability. *Transportation Research Part B* 44 (1), 38-49.
- Fosgerau, M. 2010. On the relationship between mean and variance of delay in dynamic queues with random capacity and demand. *Journal of Economic Dynamics and Control* 34 (4), 598-603.
- Hess, S., Adler, N., Polak, J.W. 2007. Modelling airport and airline choice behavior with the use of stated preference survey data. *Transportation Research Part E* 43 (3), 221-233.

- Koster, P.R., Kroes, E., Verhoef, E.T. 2011. Travel time variability and airport accessibility. *Transportation Research Part B* 45 (10), 1545-1559.
- Koster, P.R., Verhoef, E.T. 2012. A Rank-dependent Scheduling Model. *Journal of Transport Economics and Policy* 46 (1), 123-138.
- Li, Z., Hensher, D.A., Rose, J.M. 2010. Willingness to pay for travel time reliability in passenger transport: A review and some new empirical evidence. *Transportation Research Part E* 46 (3), 384-403.
- Li, Z., Tirachini, A., Hensher, D.A. 2012. Embedding risk attitudes in a scheduling model: Application to the study of commuting departure time. *Transportation Science, Articles in Advance*, 1-19.
- Lijesen, M. 2006. A mixed logit based valuation of frequency in civil aviation from SP data. *Transportation Research Part E* 42 (2), 82-94.
- Mayer, C., Sinai, T. 2003. Network effects, Congestion Externalities and Air Travel Delays: Or Why Not All Delays Are Evil. *American Economic Review* 93 (4), 1194-1215.
- Morrison, S.A., Winston, C. 2007. Another Look at Airport Congestion Pricing. *American Economic Review* 97 (5), 1970-1977.
- Noland, R.B., Polak, J.W. 2002. Travel time variability: A review of theoretical and empirical issues.
- Noland, R.B. and Small, K.A. 1995. Travel-Time uncertainty, Departure Time Choice, and the Costs of Morning Commutes, *Transportation Research Record* 1493, 150-158.
- OECD/ITF 2010. Improving reliability on surface transport networks. Downloadable via: www.internationaltransportforum.org.
- Peer, S., Verhoef, E.T., Koopmans, C. 2012. Prediction of travel time variability for cost-benefit analysis. *Transportation Research Part A* 46 (1), 79-90.
- Small, K.A. 1982. The scheduling of consumer activities: work trips. *American Economic Review* 72 (3), 469-479.
- Tseng, Y.-Y., Rietveld, P., Verhoef, E.T. 2012. Unreliable trains and induced rescheduling: implications for cost-benefit analysis. *Transportation* 39 (2), 387-407.
- van den Berg, V. A. C., Verhoef, E. T. 2011. Winning or losing from dynamic bottleneck congestion pricing?: The distributional effects of road pricing with heterogeneity in values of time and schedule delay, *Journal of Public Economics* 95 (7-8), 983-992.

Vickrey, W. 1969. Congestion theory and transport investment. *American Economic Review* 59, 251-261.

Warburg, V., Bhat, C. and Adler, T. 2006. Modeling Demographic and Unobserved Heterogeneity in Air Passengers' Sensitivity to Service Attributes in Itinerary Choice. *Transportation Research Record* 1951, 7-16.