



Munich Personal RePEc Archive

## **Modelling reliability benefits**

Eliasson, Jonas

Linköping University

March 2019

Online at <https://mpra.ub.uni-muenchen.de/94817/>  
MPRA Paper No. 94817, posted 04 Jul 2019 06:22 UTC

# Modelling reliability benefits

Jonas Eliasson  
Linköping University  
jonas.eliasson@liu.se

*Published as:*

Eliasson, J. (2019) Modeling reliability benefits. *Transport Findings*, March 2019.

Doi: [10.32866/7542](https://doi.org/10.32866/7542).

## **Abstract**

This paper compares the performance of several models forecasting travel time variability for road traffic, using before/after data from the introduction of the Stockholm congestion charges. Models are estimated on before-data, and the models' forecasts for the after-situation are compared to actual after-measurements. The accuracy of the models vary substantially, but several models are able to forecast the benefits from reduced travel time variability with sufficient accuracy to make them useful for decision making.

*Keywords:* Travel time variability, reliability, cost benefit analysis, congestion pricing.

## 1 RESEARCH QUESTIONS AND HYPOTHESES

Reduced travel time variability can constitute an important benefit when road congestion is reduced by policies or projects. A considerable literature on the valuation of reliability has emerged over the last decades, together with a smaller literature on predicting travel time variability for road traffic (de Jong and Bliemer, 2015). This paper focuses on the second issue, specifically estimating relationships between the mean and standard deviation of travel time.

Several different functional specifications have been used in the literature, and there appears to be little consensus about what works best. This paper therefore sets out to compare the performance of a number of model specifications on a particularly interesting data set: travel time measurements made before and after the introduction of congestion charges in Stockholm. The charges caused traffic to and from the inner city to decrease by around 20%, resulting in substantial congestion reductions (Eliasson, 2008). The two data sets hence represent very different traffic situations on the exact same set of links, making it ideal for evaluating and comparing the predictive performance of the models.

The fundamental question of the paper is whether models for forecasting travel time variability are sufficiently good to be useful for applied policy-making. This is explored through estimating models on data collected before the congestion charges, and then using them to forecast travel time variability in the situation with the charges. Forecasts and outcome can then be compared to see which (if any) of the different model specifications give a sufficiently correct magnitude of the reliability benefits.

## 2 METHODS AND DATA

The data consists of average travel times for 15 minute periods, collected from 41 major urban streets and arterials in and around Stockholm's inner city during six weeks in the spring of 2005 (before the charges) and the corresponding weeks in the spring of 2006 (with charges). Only data from Monday-Thursday 06:00-20:00 is used. Estimations are hence based on  $41 \cdot 14 \cdot 4 = 2296$  observations of standard deviations and average travel times, and standard deviations and average travel times for each link/time combination are calculated based on  $6 \cdot 4 = 24$  measurements each.

The links are between 0.38 and 4.82 km long, most of them with two lanes per direction, and with speed limits either 50 km/h or 70 km/h. Most of the links also include at least one intersection. Traffic volumes vary between 15 000 and 50 000 vehicles per day.

When congestion charges were introduced in Stockholm in the spring of 2006, mean travel time per kilometer (weighted by traffic volumes per link and 15 minute time period) decreased from 3.31 min/km to 2.91 min/km on the links in the sample. The corresponding mean standard deviation decreased from 1.10 min/km to 0.82 min/km. Assuming a typical reliability ratio (the ratio of the value

of standard deviation to the value of travel time) of 0.8, this indicates that reliability benefits are around 60% of conventional travel time savings. This is considerably higher than in the cost-benefit analysis reported in Eliasson (2009a), which used model calculations rather than actual measurements to estimate reliability benefits of the Stockholm charges.

Table 1 presents 9 different model specifications to be compared, drawing from the surveys in de Jong and Bliemer (2015) and Kouwenhoven and Warffemius (2016). The specifications originate from the references given, but some have been simplified to allow comparison: for example, there are no controls for e.g. road type or queueing buildup/dissipation phases, and all models are estimated using conventional linear regression, rather than more advanced error specifications (as in Kim and Mahmassani (2014)). The only variables used are mean travel time  $t$ , free-flow travel time  $t_0$ , standard deviation  $\sigma$  and link length  $d$ .

No.	Name	Formula	Reference
1	Delay polynomial	$\sigma = \beta_0 + \beta_1(t - t_0) + \beta_2(t - t_0)^2 + \beta_3(t - t_0)^3 + \beta_4d + \beta_5d^2$	(Peer et al., 2012)
2	Log-linear delay	$\sigma \sim \beta_0 + \beta_1(t - t_0) + \beta_2 \log(t - t_0 + 1) + \beta_3d$	(Kouwenhoven and Warffemius, 2016)
3	Distance-normalized linear	$\frac{\sigma}{d} \sim \beta_0 + \beta_1 \frac{t}{d}$	(Mahmassani et al., 2013)
4	Log-log time & delay	$\log(\sigma) \sim \beta_0 + \beta_1 \log(t) + \beta_2 \log\left(\frac{t}{t_0} - 1\right) + \beta_3d$	(Eliasson, 2009b)
5	Linear	$\sigma \sim \beta_0 + \beta_1t$	(Hellinga, 2011)
6	Distance-normalized log-log	$\log\left(\frac{\sigma}{d}\right) \sim \beta_0 + \beta_1 \log\left(\frac{t}{d}\right)$	(Kim and Mahmassani, 2014)
7	Square root linear	$\sigma \sim \beta_0 + \beta_1\sqrt{t}$	(Herman and Lam, 1974)
8	Time-normalized log-log	$\log\left(\frac{\sigma}{t}\right) \sim \beta_0 + \beta_1 \log\left(\frac{t}{t_0}\right) + \beta_2 * d$	(Arup et al., 2003)
9	Log-log	$\log(\sigma) \sim \beta_1 \log(t)$	(Geistefeldt et al., 2014)

Table 1. Estimated model specifications.

### 3 FINDINGS

Table 2 summarizes the essential results (full estimation results are given in Appendix). Columns 3 and 5 present  $R^2$  goodness-of-fits of the models estimated on data from each year, calculated as  $R^2_{year} = 1 - \frac{\sum_k(\hat{\sigma}_k^{year} - \sigma_k^{year})^2}{\sum_k(\sigma_k^{year} - \sigma_{mean}^{year})^2}$ , where  $\sigma_k$  are observed values and  $\hat{\sigma}_k$  are model-predicted values. The first four models (1)-(4) stand out in terms of goodness-of-fit to the estimation sample, with  $R^2$ :s around 0.5-0.6 for both samples. Columns 4 and 6 list parameter values, ignoring intercepts. Non-significant parameter values are marked “-”.

Column 7 lists  $R^2$ :s when the 2005-models are used to predict 2006 values, i.e.  $R^2 = 1 - \frac{\sum_k(\hat{\sigma}_k^{05} - \sigma_k^{06})^2}{\sum_k(\sigma_k^{06} - \sigma_{mean}^{06})^2}$ , where  $\hat{\sigma}_k^{05}$  means that the models estimated on 2005 data are used to predict observed 2006 values, using measured travel times from 2006 as input. Models (1), (2) and (4) perform remarkably well. The common factor of the best-performing models is that the standard deviation and the travel time are normalized in some sense, either by dividing by distance or by free-flow travel time, or by estimating parameter(s) for distance (which can be seen as a more flexible normalization). Of course, this also means that they have the inherent benefit of having more parameters (especially model (1) with its five parameters, some of which are not significant). The parameters of models (3)-(4) are fairly stable across years, while the parameters of models (1)-(2) vary more between years. This might be a signal of potential problems with overspecification, or be due to uncertainties in the measured variables.

Column 8 presents the most interesting result: how accurate reliability benefits are predicted. (Benefits are normalized to minutes/km, weighted by traffic volumes by link and 15-minute time period.) Again, models (1)-(4) perform well: they are close enough to the truth to pass a heuristic “usefulness” test. Models (6) and (8) also perform well – model (8) remarkably so, considering the comparatively low model fit in the estimation.

(1) No.	(2) Name	(3) Spring 2005 $R^2$	(6) Spring 2005 Parameters	(5) Spring 2006 $R^2$	(4) Spring 2006 Parameters	(7) $R^2$ of predicted 2006-values by 2005-model	(8) Over-/under-prediction of true benefits
1	Delay polynomial	0.60	$\beta_1 1.16, \beta_2 -.15, \beta_3 .01, \beta_4 -, \beta_5 -$	0.60	$\beta_1 .73, \beta_2 -, \beta_3 -, \beta_4 0.14, \beta_5 -.035$	0.57	-1%
2	Delay log-linear	0.59	$\beta_1 .36, \beta_2 .88, \beta_3 .027$	0.59	$\beta_1 .51, \beta_2 .37, \beta_3 -.041$	0.57	-3%
3	Distance-normalized linear	0.52	$\beta_1 .45$	0.46	$\beta_1 .49$	0.39	-19%
4	Time&delay log-log	0.48	$\beta_1 1.07, \beta_2 .52, \beta_3 -.15$	0.53	$\beta_1 1.06, \beta_2 .49, \beta_3 -.18$	0.56	17%
5	Linear	0.39	$\beta_1 .35$	0.25	$\beta_1 .25$	0.17	51%
6	Distance-normalized log-log	0.35	$\beta_1 1.79$	0.42	$\beta_1 1.81$	0.46	27%
7	Square root linear	0.34	$\beta_1 1.36$	0.22	$\beta_1 0.92$	0.12	62%
8	Time-normalized log-log	0.33	$\beta_1 1.30, \beta_2 -.14$	0.24	$\beta_1 1.60, \beta_2 -.13$	0.43	7%
9	Log-log	0.26	$\beta_1 1.06$	0.12	$\beta_1 0.90$	0.24	95%

Table 2. Estimation results. Explanations in the text.

Results indicate that it is indeed possible to predict reliability benefits with an accuracy that is sufficiently good to be useful for policy making. Note, though, that

travel times need to be predicted well for this to work, since travel times are essential inputs to the variability models: the comparisons in this paper use actual travel time measurements from the prediction year (2006), in order not to confound the performance of the variability forecasting model (which is the focus of this paper) with the performance of a travel time forecasting model.

There are substantial differences between the models' performances, however. Having enough flexibility in the model seems to pay off, giving remarkably good predictions of total benefits. The slight instability of parameters across different samples does not seem to be a real problem, since the out-of-sample predictions are so good. Of the simpler models, the distance-normalized linear model stands out for performing well with only two parameters.

## 4 REFERENCES

- Arup, Bates, J., Fearon, J., Black, I., 2003. Frameworks for Modelling the Variability of Journey Times on the Highway Network. Arup, London.
- de Jong, G.C., Bliemer, M.C.J., 2015. On including travel time reliability of road traffic in appraisal. *Transportation Research Part A: Policy and Practice* 73, 80–95. <https://doi.org/10.1016/j.tra.2015.01.006>
- Eliasson, J., 2009a. A cost-benefit analysis of the Stockholm congestion charging system. *Transportation Research Part A: Policy and Practice* 43, 468–480. <https://doi.org/10.1016/j.tra.2008.11.014>
- Eliasson, J., 2009b. Forecasting travel time variability, in: *Proceedings of the European Transport Conference*.
- Eliasson, J., 2008. Lessons from the Stockholm congestion charging trial. *Transport Policy* 15, 395–404. <https://doi.org/10.1016/j.tranpol.2008.12.004>
- Geistefeldt, J., Hohmann, S., Wu, N., 2014. Ermittlung des Zusammenhangs von Infrastruktur und Zuverlässigkeit des Verkehrsablaufs für den Verkehrsträger Strasse. Ruhr Universität Bochum.
- Hellinga, B., 2011. Defining, measuring, and modeling transportation network reliability. Dutch Ministry of Transport, Rijkswaterstaat, Delft.
- Herman, R., Lam, T.C., 1974. Trip Time Characteristics of Journeys to and from Work, in: *Proceedings. Presented at the Transportation and Traffic Theory*.
- Kim, J., Mahmassani, H.S., 2014. A finite mixture model of vehicle-to-vehicle and day-to-day variability of traffic network travel times. *Transportation Research Part C: Emerging Technologies* 46, 83–97. <https://doi.org/10.1016/j.trc.2014.05.011>
- Kouwenhoven, M., Warffemius, P., 2016. Forecasting travel-time reliability in road transport: A new model for the Netherlands (Discussion Paper No. 2016– 02), ITF Roundtable Reports. International Transport Forum.
- Mahmassani, H.S., Hou, T., Saberi, M., 2013. Connecting Networkwide Travel Time Reliability and the Network Fundamental Diagram of Traffic Flow. *Transportation Research Record* 2391, 80–91. <https://doi.org/10.3141/2391-08>

## 5 APPENDIX: ESTIMATION RESULTS

\*\*\*\*\*  
 Models estimated on spring 2005 data  
 \*\*\*\*\*

Call:  
 lm(formula = as.formula(KouwForm), data = datauttag)

Residuals:  
 Min 1Q Median 3Q Max  
 -2.4760 -0.3731 -0.1131 0.1783 4.7344

Coefficients:  

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-0.22018	0.06084	-3.619	0.000304 ***
I(restid - fftid)	0.36251	0.04296	8.439	< 2e-16 ***
log(restid - fftid + 1)	0.88242	0.13594	6.491	1.08e-10 ***
langd	0.02683	0.01266	2.119	0.034253 *

---  
 Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7703 on 1951 degrees of freedom  
 Multiple R-squared: 0.5895, Adjusted R-squared: 0.5889  
 F-statistic: 934.1 on 3 and 1951 DF, p-value: < 2.2e-16

\*\*\*\*\*

Call:  
 lm(formula = as.formula(MahmlForm), data = datauttag)

Residuals:  
 Min 1Q Median 3Q Max  
 -2.6259 -0.2222 -0.0861 0.1087 6.1732

Coefficients:  

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-0.298356	0.023821	-12.53	<2e-16 ***
I(restid/langd)	0.452338	0.008845	51.14	<2e-16 ***

---  
 Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5854 on 1953 degrees of freedom  
 Multiple R-squared: 0.5725, Adjusted R-squared: 0.5723  
 F-statistic: 2615 on 1 and 1953 DF, p-value: < 2.2e-16

\*\*\*\*\*

Call:  
 lm(formula = as.formula(EliasForm), data = datauttag)

Residuals:  
 Min 1Q Median 3Q Max  
 -2.02707 -0.40941 -0.05681 0.34013 2.25230

Coefficients:  

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-1.25360	0.04666	-26.868	< 2e-16 ***
log(restid)	1.07109	0.03967	26.999	< 2e-16 ***
log(rel.fordroj)	0.52528	0.01955	26.865	< 2e-16 ***

```

log(langd)      -0.15237    0.03343  -4.558 5.47e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5696 on 1951 degrees of freedom
Multiple R-squared:  0.6647,    Adjusted R-squared:  0.6642
F-statistic: 1289 on 3 and 1951 DF,  p-value: < 2.2e-16

*****

Call:
lm(formula = as.formula(HellingaForm), data = dataauttag)

Residuals:
    Min       1Q   Median       3Q      Max
-2.1091 -0.6158 -0.1234  0.4023  6.0817

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.328930   0.047114  -6.982 3.99e-12 ***
restid       0.347681   0.009843  35.321 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9388 on 1953 degrees of freedom
Multiple R-squared:  0.3898,    Adjusted R-squared:  0.3895
F-statistic: 1248 on 1 and 1953 DF,  p-value: < 2.2e-16

*****

Call:
lm(formula = as.formula(Mahm2Form), data = dataauttag)

Residuals:
    Min       1Q   Median       3Q      Max
-1.86663 -0.49532 -0.05867  0.44919  2.41020

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -2.14375   0.02285  -93.82  <2e-16 ***
log(restid/langd)  1.78844   0.02699   66.26  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6672 on 1953 degrees of freedom
Multiple R-squared:  0.6921,    Adjusted R-squared:  0.692
F-statistic: 4391 on 1 and 1953 DF,  p-value: < 2.2e-16

*****

Call:
lm(formula = as.formula(HermanForm), data = dataauttag)

Residuals:
    Min       1Q   Median       3Q      Max
-1.8685 -0.6562 -0.1372  0.4299  7.2340

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.57442   0.08818  -17.86  <2e-16 ***
sqrt(restid)  1.36444   0.04266   31.98  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9735 on 1953 degrees of freedom
Multiple R-squared:  0.3438,    Adjusted R-squared:  0.3434
F-statistic: 1023 on 1 and 1953 DF,  p-value: < 2.2e-16

```



```

*****
Call:
lm(formula = as.formula(ArupForm), data = datauttag)

Residuals:
    Min       1Q   Median       3Q      Max
-1.86293 -0.44557 -0.07692  0.38285  2.26744

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   -2.18263    0.03368  -64.811 < 2e-16 ***
log(restid/fftid)  1.29527    0.04295   30.159 < 2e-16 ***
log(langd)     -0.14117    0.02146   -6.577 6.14e-11 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6059 on 1952 degrees of freedom
Multiple R-squared:  0.4263,    Adjusted R-squared:  0.4257
F-statistic: 725.1 on 2 and 1952 DF,  p-value: < 2.2e-16

```

```

*****
Call:
lm(formula = as.formula(BMVDIForm), data = datauttag)

Residuals:
    Min       1Q   Median       3Q      Max
-2.06883 -0.66164  0.01742  0.63670  2.32833

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -1.71733    0.04753  -36.13 <2e-16 ***
log(restid)  1.05678    0.03336   31.68 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7992 on 1953 degrees of freedom
Multiple R-squared:  0.3394,    Adjusted R-squared:  0.339
F-statistic: 1003 on 1 and 1953 DF,  p-value: < 2.2e-16

```

```

*****
Call:
lm(formula = as.formula(PeerForm), data = datauttag)

Residuals:
    Min       1Q   Median       3Q      Max
-3.1176 -0.3524 -0.1420  0.1554  4.8191

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.02655    0.04103    0.647 0.517710
I(restid - fftid) 0.48849    0.02855  17.110 < 2e-16 ***
restid          0.12935    0.02362    5.477 4.89e-08 ***
langd          -0.08072    0.02412   -3.347 0.000833 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7727 on 1951 degrees of freedom
Multiple R-squared:  0.587,    Adjusted R-squared:  0.5864
F-statistic: 924.4 on 3 and 1951 DF,  p-value: < 2.2e-16

```

```

*****
*****
Models estimated on spring 2006 data
*****

```

```

Call:
lm(formula = as.formula(KouwForm), data = datauttag)

Residuals:
    Min       1Q   Median       3Q      Max
-2.2407 -0.2824 -0.1154  0.1316  5.2774

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.04374    0.04525   0.966  0.33393
I(restid - fftid)  0.51253    0.04017  12.759 < 2e-16 ***
log(restid - fftid + 1) 0.36922    0.11485   3.215  0.00133 **
langd          -0.04137    0.01022  -4.050  5.34e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6119 on 1908 degrees of freedom
Multiple R-squared:  0.5913,    Adjusted R-squared:  0.5906
F-statistic: 920.1 on 3 and 1908 DF,  p-value: < 2.2e-16

```

\*\*\*\*\*

```

Call:
lm(formula = as.formula(MahmlForm), data = datauttag)

Residuals:
    Min       1Q   Median       3Q      Max
-2.1828 -0.1807 -0.0207  0.0931  4.3258

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.454966    0.018280  -24.89 <2e-16 ***
I(restid/langd)  0.491873    0.007932   62.01 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4214 on 1910 degrees of freedom
Multiple R-squared:  0.6682,    Adjusted R-squared:  0.668
F-statistic: 3846 on 1 and 1910 DF,  p-value: < 2.2e-16

```

\*\*\*\*\*

```

Call:
lm(formula = as.formula(EliasForm), data = datauttag)

Residuals:
    Min       1Q   Median       3Q      Max
-1.6089 -0.4298 -0.0689  0.3261  2.5156

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  -1.36607    0.04728  -28.893 < 2e-16 ***
log(restid)    1.06441    0.04351  24.465 < 2e-16 ***
log(rel.fordroj) 0.49302    0.01786  27.608 < 2e-16 ***
log(langd)    -0.17831    0.03602  -4.951 8.04e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6039 on 1908 degrees of freedom
Multiple R-squared:  0.5983,    Adjusted R-squared:  0.5977
F-statistic: 947.3 on 3 and 1908 DF,  p-value: < 2.2e-16

```

\*\*\*\*\*

```

Call:
lm(formula = as.formula(HellingaForm), data = datauttag)

Residuals:

```

```
      Min      1Q  Median      3Q      Max
-1.1882 -0.5500 -0.1550  0.2466  6.4721
```

Coefficients:

```
      Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.133776   0.042161  -3.173  0.00153 **
restid       0.246342   0.009671  25.474 < 2e-16 ***
```

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8264 on 1910 degrees of freedom  
Multiple R-squared: 0.2536, Adjusted R-squared: 0.2532  
F-statistic: 648.9 on 1 and 1910 DF, p-value: < 2.2e-16

\*\*\*\*\*

Call:

```
lm(formula = as.formula(Mahm2Form), data = dataouttag)
```

Residuals:

```
      Min      1Q  Median      3Q      Max
-2.19061 -0.45885 -0.09676  0.39846  2.53083
```

Coefficients:

```
      Estimate Std. Error t value Pr(>|t|)
(Intercept) -2.32870   0.02399  -97.06 <2e-16 ***
log(restid/langd) 1.81041   0.03285   55.11 <2e-16 ***
```

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7173 on 1910 degrees of freedom  
Multiple R-squared: 0.6139, Adjusted R-squared: 0.6137  
F-statistic: 3037 on 1 and 1910 DF, p-value: < 2.2e-16

\*\*\*\*\*

Call:

```
lm(formula = as.formula(HermanForm), data = dataouttag)
```

Residuals:

```
      Min      1Q  Median      3Q      Max
-1.0949 -0.5813 -0.1577  0.2881  6.7074
```

Coefficients:

```
      Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.93653   0.07762  -12.06 <2e-16 ***
sqrt(restid) 0.92183   0.03932   23.44 <2e-16 ***
```

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8429 on 1910 degrees of freedom  
Multiple R-squared: 0.2235, Adjusted R-squared: 0.2231  
F-statistic: 549.6 on 1 and 1910 DF, p-value: < 2.2e-16

\*\*\*\*\*

Call:

```
lm(formula = as.formula(ArupForm), data = dataouttag)
```

Residuals:

```
      Min      1Q  Median      3Q      Max
-1.58053 -0.43557 -0.08308  0.31956  2.05701
```

Coefficients:

```
      Estimate Std. Error t value Pr(>|t|)
(Intercept) -2.46245   0.03345 -73.618 < 2e-16 ***
log(restid/fftid) 1.60613   0.05001  32.119 < 2e-16 ***
log(langd)    -0.12735   0.02176  -5.852 5.72e-09 ***
```

```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6112 on 1909 degrees of freedom
Multiple R-squared:  0.4498,    Adjusted R-squared:  0.4492
F-statistic: 780.4 on 2 and 1909 DF,  p-value: < 2.2e-16

```

\*\*\*\*\*

```

Call:
lm(formula = as.formula(BMVDIForm), data = datauttag)

```

```

Residuals:
    Min       1Q   Median       3Q      Max
-1.69470 -0.70306 -0.08463  0.65004  2.32607

```

```

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.76975    0.04692  -37.72  <2e-16 ***
log(restid)  0.89644    0.03503   25.59  <2e-16 ***
---

```

```

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

Residual standard error: 0.8218 on 1910 degrees of freedom
Multiple R-squared:  0.2553,    Adjusted R-squared:  0.2549
F-statistic: 654.9 on 1 and 1910 DF,  p-value: < 2.2e-16

```

\*\*\*\*\*

```

Call:
lm(formula = as.formula(PeerForm), data = datauttag)

```

```

Residuals:
    Min       1Q   Median       3Q      Max
-2.3818 -0.2737 -0.1162  0.1005  5.2276

```

```

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.13160    0.03242   4.059 5.12e-05 ***
I(restid - fftid) 0.58759    0.02434  24.140 < 2e-16 ***
restid          0.04451    0.01954   2.278  0.0229 *
langd           -0.07895    0.02019  -3.911 9.53e-05 ***
---

```

```

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

Residual standard error: 0.6127 on 1908 degrees of freedom
Multiple R-squared:  0.5902,    Adjusted R-squared:  0.5895
F-statistic: 915.9 on 3 and 1908 DF,  p-value: < 2.2e-16

```

\*\*\*\*\*