Postal Rate Commission
Submitted 5/26/2005 4:13 pm
Filing ID: 44451
Accepted 5/26/2005

BEFORE THE
POSTAL RATE COMMISSION
WASHINGTON, D.C. 20268–0001

POSTAL RATE AND FEE CHANGES
PURSUANT TO PUBLIC LAW 108-18

Docket No. R2005–1

RESPONSES OF POSTAL SERVICE WITNESS BRADLEY
TO INTERROGATORIES OF ADVO (ADVO/USPS-T14-3 - 15)
(May 26, 2005)

The United States Postal Service hereby provides the responses of witness Bradley to the following interrogatories of Advo: Advo/USPS-T14-3 – 14, filed on May 12, and Advo/USPS-T14-15, filed on May 16, 2005.

Each interrogatory is stated verbatim and is followed by the response.

Respectfully submitted,

UNITED STATES POSTAL SERVICE

By its attorneys:

Daniel J. Foucheaux, Jr.
Chief Counsel, Ratemaking

____
Eric P. Koetting

475 L'Enfant Plaza West, S.W.
Washington, D.C. 20260-1137
(202) 268-2992, Fax -5402
May 26, 2005

ADVO/USPS-T14-3. On page 33 of your testimony, you state: "Because of the large cross-sectional variation in the data, it is likely that econometric estimates for the delivery equations suffer from heteroskedasticity."

(a) Do you mean that cross-sectional data normally exhibit characteristics that cause error variances to change in size with variations in one or more of the independent variables? Please explain fully.

(b) Did you conduct any diagnostic tests to detect heteroskedasticity in your recommended and alternative models? If so, please provide results from these tests. If not, please explain why these tests were not conducted.

(c) Please confirm that use of ordinary least squares when heteroskedasticity is present leads to unbiased but inefficient parameter estimates. If not, please explain fully.

ADVO/USPS-T14-3. Response:

a. I was not suggesting that it is normal for the error variances in a cross sectional regression to be heteroskedastic, but rather that presence of heteroskedasticity is a common problem in cross-sectional regressions.

b. No. Because I intended to correct for heteroskedasticity using White's HCSE approach, I did not first test for its presence. Note that the HCSE approach does not alter the point estimates, so it is not imprudent to go ahead and perform the correction and then check the results.

c. Confirmed.

ADVO/USPS-T14-4. On page 52 of your testimony, you state that the "actual form of heteroskedasticity is unknown".

(a) Please provide and explain in general terms alternatives to ordinary least squares that correct for heteroskedastic data, when the actual form of heteroskedasticity is known, and what form of heteroskedasticity each corrects.

(b) When the actual form of heteroskedasticity is unknown, as you state, is it accepted econometric procedure to infer possible causes, apply the corresponding corrective procedure and then test for heteroskedasticity ex-post to determine whether the initial inference was correct?

(c) Please note that on page 52 you also describe certain assumptions leading to a weighted least squares procedure. Please confirm that if the standard error of city carrier street time cost is suspected to be correlated with zip-code area size, then it would be appropriate to run a weighted least squares regression using the inverse of area square mileage as the weighting factor. If you cannot confirm, please explain why.

(d) If you do confirm in (c) above, then is it appropriate to determine whether or not the transformed error term (through the weighted least squares procedure) is now homoskedastic through appropriate diagnostic testing? Please explain fully.

(e) Are there circumstances when the heteroskedasticity form can be known a-priori and the appropriate corrective procedure applied without further diagnostic testing? If so what are these circumstances? Please explain fully.

ADVO/USPS-T14-4 Response:

a. If the actual form of the heteroskedasticity is known, then generalized least squares (GLS) can be used to estimate the equation. Recall that under homoskedastic errors, the error structure is given by $\sigma^2\mathbf{I}$, where **I** is an identity matrix. When the form of heteroskedasticity is known, then the error structure is given by $\sigma^2\boldsymbol{\Omega}$, where $\boldsymbol{\Omega}$ is a matrix which embodies the specification of the error variance. Because **Ω** is known, the model can be estimated using the GLS estimator: [1]

$$\tilde{\beta} = \left[X'\Omega^{-1}X\right]^{-1} X'\Omega^{-1} y .$$

b. When the form of the heteroskedasticiy is unknown, one of two approaches can be followed. First, one could apply feasible generalized least squares (FGLS) in which consistent estimates of elements **Ω** are first obtained and then the GLS estimator presented above is used with an estimated **Ω** in place of the actual one:

$$\tilde{\tilde{\beta}} = \left[X'\hat{\Omega}^{-1}X\right]^{-1} X'\hat{\Omega}^{-1} y$$

Second, White's estimator can be used to estimate a heteroskedasticity robust variance/covariance matrix:[2]

> It may well be that the form of heteroscedasticity is unknown. White (1980) has shown that it is still possible to obtain an appropriate estimator for the variance of the least squares estimator, even if the heteroscedasticity is related to the variables in **X.**

---

1 Please note that weighted least squares is one form of GLS.

2 See, Greene, William H., Econometric Analysis, Macmillian Publishing Co., NY, 1993, at 391.

White's estimator provides a consistent estimator of the variance/covariance matrix for the regression and permits calculation of consistent estimates of standard errors. This yields unbiased statistical inferences about the parameters:[3]

> This is an extremely important and useful result. It implies that without actually specifying the type of heteroscedasticity, we can still make appropriate inferences based on the results of least squares. This is especially useful if we are unsure of the precise nature of the heteroscedasticity (which is probably most of the time).

c. Not confirmed. A weighted least squares approach is appropriate when the form of the heteroskedasticity is known. This requires additional knowledge or a stronger assumption than just the fact that the standard error of the regression is correlated with an auxiliary variable. One needs to specify the skedasticity function before estimation of the model. That is, the nature of the relationship between the standard error and the auxiliary variable must be specified. In addition, when the nature of the heteroskedasticity is not known, White's estimator provides an approach that permits estimation of a heteroskedastic consistent covariance matrix. This approach is thus widely used.

d. Part c was not confirmed. Generally, the error variances are not tested after correction.

---

[3] Id.

e. If the analyst has extra-sample information about the nature of the heteroskedasticity, then that information can be used to specify the variance covariance matrix for GLS estimator.

ADVO/USPS-T14-5. Please assume a simple zip-code delivery cost model stated in functional form for N number of zip-codes as:

$C_i = f(V_i, PD_i, A_i) + e_i$,

where: $i = 1, 2, \ldots, N$ and,

C_i = zip-code i delivery cost,
V_i = zip-code i volume,
PD_i = the number of zip-code i possible deliveries,
A_i = the zip-code i total area in square miles.

Further the function is assumed to be homogenous to the first degree so that:

$C_i{*}k = f(V_i{*}k, PD_i{*}k, A_i{*}k) + e_i{*}k$.

(a) Please confirm that such a model would predict positive marginal cost effects with respect to the three workload variables independently. If you cannot confirm, please explain why not.

(b) Please confirm that in such a model, if values for the three explanatory variables were twice as high in one zip-code compared to another, the model would predict total delivery costs that were also twice as high in the former zip-code compared to the latter. If not, please explain why not.

(c) Please confirm that such a model would predict a volume variability less than one. If not, please explain why not.

(d) Please comment on the general characterization of delivery costs that are assumed to behave as described by the model. Please explain fully under what conditions such a behavioral structure might be expected. Alternately, might one expect systematic deviations from model predictions as zip-code square miles increase and all other variables grow in the same proportion? If so, please explain fully.

ADVO/USPS-T14-5 Response:

Please note that the question is silent relative to defining the variable e_i. However, the logic of parts a. through c. depends upon the value of e_i being equal to zero. I believe that you meant e_i to be a regression error term with an expected value of zero. Thus, I will treat e_i as having a value of zero in parts a through c.

a. Not confirmed by contradiction. Consider the following function (note I suppress the subscript for notational convenience):

$C = V^{0.4}\ PD^{0.8}\ A^{-0.2}$.

First, check the homogeneity of the function:

$\lambda^k C = (\lambda V)^{0.4}\ (\lambda\ PD)^{0.8}\ (\lambda A)^{-0.2}$.

so:

$\lambda^k C = \lambda^{0.4} \lambda^{0.8} \lambda^{-0.2}\ V^{0.4}\ PD^{0.8}\ A^{-0.2}$.

or:

$\lambda^k C = \lambda^{(0.4+0.8-0.2)}\ (V^{0.4}\ PD^{0.8}\ A^{-0.2})$.

so:

$\lambda^k C = \lambda C$. Thus, k =1 and the function is homogenous of degree one. Now, calculate the marginal cost of A:

$$MC_A = \frac{\partial C}{\partial A} = -1.2 V^{0.4} PD^{0.8} A^{-1.2} < 0.$$

b. Confirmed

c. Not confirmed by contradiction. Consider the function:

$C = V^{1.2}\ PD^{0.8}\ A^{-1.0}$.

This function is homogenous of degree one but has a volume variability greater than one. Please note that if one first restricts all parameters to be greater than zero in this function, then the assumption of first degree homogeneity ensures that the coefficient on "V" (the volume variability) is less than one.

d. As I did not create the model, it is difficult for me to infer the underlying behavioral assumptions used to generate it. If one is assuming the delivery costs are generated by a function that is homogeneous of degree one in the three specified variables, then one is starting from the *a priori* belief that delivery is "replicable" in the sense that if all three variables are doubled then the cost of delivery is doubled. In response to the second part of the question, if the assumed model is correct, then one would not expect systematic deviations from model predictions as zip-code square miles increase and all other variables grow in the same proportion.

ADVO/USPS-T14-6. Please provide correlation matrices for all regression results presented in your testimony, including all alternative models you present but do not recommend for development of volume variable costs.

ADVO/USPS-T14-6 Response:

Below are the correlation matrixes for the three models I present in my testimony, the full quadratic regular delivery equation, the restricted quadratic regular delivery equation, and the parcel/accountable equation. I did not produce correlation matrices for the alternative models estimated along the research path as they do not produce additional information that is required for model evaluation.

Correlation Matrix for the Full Quadratic Regular Delivery Equation

Correlation of Estimates

Variable	Intercept	let	let2	cf	cf2
Intercept	1.0000	0.0705	0.0201	-0.0245	0.0530
let	0.0705	1.0000	0.1043	-0.5443	0.1953
let2	0.0201	0.1043	1.0000	-0.0631	0.1675
cf	-0.0245	-0.5443	-0.0631	1.0000	-0.4728
cf2	0.0530	0.1953	0.1675	-0.4728	1.0000
seq	-0.0806	-0.0473	-0.0188	0.0266	-0.0031
seq2	0.0085	0.0131	-0.0203	0.0336	-0.0478
cv	-0.2685	-0.0430	0.0831	-0.1562	0.1216
cv2	-0.0096	0.0078	-0.0177	0.0423	-0.1477
spr	-0.0564	-0.2914	-0.0475	-0.0838	-0.0393
spr2	0.0609	0.0192	0.2203	-0.0723	0.0606
dp	-0.4470	-0.4616	-0.0857	0.0287	0.0324
dp2	0.1540	0.2891	0.2479	-0.1166	0.0597
dens	-0.3754	-0.0956	0.0542	-0.1177	-0.0122
dens2	0.2421	0.0757	-0.0481	0.1494	-0.0100
lf	-0.0133	-0.1942	-0.6692	0.1496	-0.4747
lse	0.0017	-0.2645	0.0676	0.1345	0.0218
lcv	-0.0401	-0.2709	-0.0006	0.1404	0.0655
lspr	-0.0288	-0.0889	-0.5304	0.0528	0.0023
ldp	-0.0257	-0.4425	-0.5161	0.2206	-0.0290
fse	0.0061	0.1659	-0.0319	-0.2844	0.0638
fcv	-0.0250	0.1191	-0.0648	-0.1731	-0.1474
fspr	-0.0249	0.1222	0.3939	-0.0631	-0.0370
fdp	0.0075	0.2306	0.3219	-0.2667	-0.1189
scv	0.0094	-0.0198	-0.0487	0.0160	-0.0441
sspr	0.0195	0.1110	0.0690	-0.0451	0.0169
sdp	0.0230	0.1129	-0.0444	0.0056	-0.0073
cspr	0.0250	0.1091	0.0051	-0.0209	-0.0466
cdp	0.2493	0.1822	0.0092	0.0453	0.0307

spdp	0.0251	0.1298	0.0628	0.1081	-0.0059
ldns	-0.0567	-0.2820	-0.4345	0.1865	-0.1530
fdns	-0.0061	0.1877	0.1863	-0.4408	0.2231
sdns	0.1229	0.0313	-0.0140	0.0072	-0.0236
cdns	0.2505	0.0478	-0.0773	0.1321	-0.0431
spdns	0.0651	0.0976	0.1360	0.0570	0.0907
dpdns	0.1530	0.0727	0.1022	0.0938	-0.0438

Correlation of Estimates

Variable	seq	seq2	cv	cv2	spr
Intercept	-0.0806	0.0085	-0.2685	-0.0096	-0.0564
let	-0.0473	0.0131	-0.0430	0.0078	-0.2914
let2	-0.0188	-0.0203	0.0831	-0.0177	-0.0475
cf	0.0266	0.0336	-0.1562	0.0423	-0.0838
cf2	-0.0031	-0.0478	0.1216	-0.1477	-0.0393
seq	1.0000	-0.3217	-0.1307	0.0137	0.0339
seq2	-0.3217	1.0000	0.0421	0.1870	0.0629
cv	-0.1307	0.0421	1.0000	-0.2858	-0.0234
cv2	0.0137	0.1870	-0.2858	1.0000	0.0813
spr	0.0339	0.0629	-0.0234	0.0813	1.0000
spr2	0.0144	-0.0530	0.0610	-0.0999	0.0904
dp	-0.0977	-0.0434	-0.0062	0.0482	-0.2847
dp2	0.0182	-0.0029	0.0297	-0.0065	0.1684
dens	0.0951	0.0208	0.2236	0.0629	0.1686
dens2	-0.0444	-0.0201	-0.1299	-0.0480	-0.1424
lf	0.0193	-0.0008	-0.0728	-0.0156	0.0856
lse	0.0268	-0.0049	-0.0368	-0.0453	0.0840
lcv	-0.0020	0.0750	0.0386	0.0956	0.0593
lspr	-0.0088	0.0026	-0.0675	0.0483	0.0524
ldp	0.0579	0.0067	-0.0160	-0.0497	0.1292
fse	-0.3083	0.0273	0.0713	0.0396	-0.0388
fcv	-0.0407	-0.0522	-0.1689	0.2680	-0.0145
fspr	-0.0349	0.0196	0.0151	0.0237	-0.2026
fdp	0.0137	-0.0036	0.1084	0.0464	0.1661
scv	-0.1431	-0.2946	-0.1987	-0.3949	-0.0142
sspr	-0.0165	-0.2583	0.0371	0.0650	-0.2576
sdp	-0.3304	0.1176	0.1425	0.0798	0.0871
cspr	-0.0392	-0.1692	-0.0748	-0.3418	-0.2460
cdp	0.1743	0.0369	-0.4363	-0.0859	0.0369
spdp	-0.0131	0.0206	0.0116	0.0322	-0.5071
ldns	-0.0513	0.0027	0.1351	0.1060	0.0602
fdns	-0.0147	0.0185	-0.0074	0.0079	0.0510
sdns	-0.4597	0.0967	0.0506	-0.0804	-0.0062
cdns	0.1022	-0.1271	-0.7863	-0.1945	-0.0307
spdns	0.0471	-0.0282	-0.0294	-0.1024	-0.3615
dpdns	0.0004	-0.0046	-0.1289	-0.0647	0.0908

Correlation of Estimates

Variable	spr2	dp	dp2	dens	dens2
Intercept	0.0609	-0.4470	0.1540	-0.3754	0.2421
let	0.0192	-0.4616	0.2891	-0.0956	0.0757
let2	0.2203	-0.0857	0.2479	0.0542	-0.0481
cf	-0.0723	0.0287	-0.1166	-0.1177	0.1494
cf2	0.0606	0.0324	0.0597	-0.0122	-0.0100
seq	0.0144	-0.0977	0.0182	0.0951	-0.0444
seq2	-0.0530	-0.0434	-0.0029	0.0208	-0.0201
cv	0.0610	-0.0062	0.0297	0.2236	-0.1299
cv2	-0.0999	0.0482	-0.0065	0.0629	-0.0480
spr	0.0904	-0.2847	0.1684	0.1686	-0.1424
spr2	1.0000	-0.1062	0.1673	-0.0637	0.0774
dp	-0.1062	1.0000	-0.5012	-0.0062	0.0046

dp2	0.1673	-0.5012	1.0000	0.0239	0.0444
dens	-0.0637	-0.0062	0.0239	1.0000	-0.8708
dens2	0.0774	0.0046	0.0444	-0.8708	1.0000
lf	-0.0819	0.0688	-0.0576	-0.0100	0.0202
lse	-0.0400	0.1185	0.1018	-0.0049	0.0031
lcv	-0.0032	0.1464	-0.0358	0.0699	-0.0536
lspr	-0.4633	0.0741	0.1899	-0.0228	0.0271
ldp	0.0761	0.2359	-0.6524	-0.0138	-0.0012
fse	0.0117	0.0530	-0.0526	0.0382	-0.0462
fcv	-0.0214	0.1134	-0.0683	0.0083	-0.0349
fspr	-0.0110	0.0777	0.0425	0.0053	-0.0195
fdp	0.0702	-0.1978	0.0867	0.1012	-0.0801
scv	-0.0446	0.1423	-0.0946	-0.0758	0.0616
sspr	0.1629	0.0621	0.0491	-0.0284	0.0342
sdp	-0.0431	-0.1574	-0.1024	0.0136	-0.0319
cspr	-0.1660	0.1018	-0.0733	-0.0809	0.0711
cdp	0.1168	-0.3260	0.1315	-0.1690	0.1019
spdp	-0.4138	0.1097	-0.4622	0.0069	-0.0452
ldns	-0.2466	0.1346	0.1366	0.0038	0.0499
fdns	0.0733	0.0380	0.0534	0.0330	-0.1523
sdns	-0.0396	-0.0241	0.0699	-0.2984	0.2077
cdns	-0.0088	0.0073	-0.0619	-0.3839	0.2548
spdns	0.2928	0.0911	0.1063	-0.2565	0.2807
dpdns	-0.0465	-0.2010	-0.3828	-0.1245	-0.0308

Correlation of Estimates

Variable	lf	lse	lcv	lspr	ldp
Intercept	-0.0133	0.0017	-0.0401	-0.0288	-0.0257
let	-0.1942	-0.2645	-0.2709	-0.0889	-0.4425
let2	-0.6692	0.0676	-0.0006	-0.5304	-0.5161
cf	0.1496	0.1345	0.1404	0.0528	0.2206
cf2	-0.4747	0.0218	0.0655	0.0023	-0.0290
seq	0.0193	0.0268	-0.0020	-0.0088	0.0579
seq2	-0.0008	-0.0049	0.0750	0.0026	0.0067
cv	-0.0728	-0.0368	0.0386	-0.0675	-0.0160
cv2	-0.0156	-0.0453	0.0956	0.0483	-0.0497
spr	0.0856	0.0840	0.0593	0.0524	0.1292
spr2	-0.0819	-0.0400	-0.0032	-0.4633	0.0761
dp	0.0688	0.1185	0.1464	0.0741	0.2359
dp2	-0.0576	0.1018	-0.0358	0.1899	-0.6524
dens	-0.0100	-0.0049	0.0699	-0.0228	-0.0138
dens2	0.0202	0.0031	-0.0536	0.0271	-0.0012
lf	1.0000	-0.0010	-0.2360	0.2108	0.1956
lse	-0.0010	1.0000	-0.0889	0.1471	-0.1635
lcv	-0.2360	-0.0889	1.0000	0.0779	0.0620
lspr	0.2108	0.1471	0.0779	1.0000	-0.1552
ldp	0.1956	-0.1635	0.0620	-0.1552	1.0000
fse	0.0001	-0.5237	0.0726	-0.0783	0.0545
fcv	0.1021	0.0437	-0.5419	-0.0078	0.0301
fspr	-0.4000	-0.1062	0.0393	-0.4926	0.0611
fdp	-0.4075	0.0423	0.1338	0.1328	-0.4501
scv	0.0764	0.0325	-0.1494	-0.0219	0.0713
sspr	-0.0646	-0.3318	-0.0106	-0.0626	-0.0153
sdp	-0.0048	-0.4734	0.0881	-0.0416	0.0574
cspr	0.1335	0.0350	-0.5589	-0.0401	-0.0118
cdp	0.1311	0.0859	-0.4459	-0.0242	-0.0627
spdp	0.0823	-0.0348	-0.0559	-0.2682	-0.0424
ldns	0.3717	0.2336	-0.0306	0.5631	-0.0962
fdns	-0.2293	-0.1163	0.0393	-0.2466	0.0392
sdns	-0.0009	0.0137	0.0055	0.0347	-0.0429
cdns	0.0919	0.0539	-0.1216	0.0282	0.0530
spdns	-0.1529	-0.0439	-0.0030	-0.1061	-0.0220
dpdns	-0.0389	-0.0764	-0.0668	-0.2278	0.1192

Correlation of Estimates

Variable	fse	fcv	fspr	fdp	scv
Intercept	0.0061	-0.0250	-0.0249	0.0075	0.0094
let	0.1659	0.1191	0.1222	0.2306	-0.0198
let2	-0.0319	-0.0648	0.3939	0.3219	-0.0487
cf	-0.2844	-0.1731	-0.0631	-0.2667	0.0160
cf2	0.0638	-0.1474	-0.0370	-0.1189	-0.0441
seq	-0.3083	-0.0407	-0.0349	0.0137	-0.1431
seq2	0.0273	-0.0522	0.0196	-0.0036	-0.2946
cv	0.0713	-0.1689	0.0151	0.1084	-0.1987
cv2	0.0396	0.2680	0.0237	0.0464	-0.3949
spr	-0.0388	-0.0145	-0.2026	0.1661	-0.0142
spr2	0.0117	-0.0214	-0.0110	0.0702	-0.0446
dp	0.0530	0.1134	0.0777	-0.1978	0.1423
dp2	-0.0526	-0.0683	0.0425	0.0867	-0.0946
dens	0.0382	0.0083	0.0053	0.1012	-0.0758
dens2	-0.0462	-0.0349	-0.0195	-0.0801	0.0616
lf	0.0001	0.1021	-0.4000	-0.4075	0.0764
lse	-0.5237	0.0437	-0.1062	0.0423	0.0325
lcv	0.0726	-0.5419	0.0393	0.1338	-0.1494
lspr	-0.0783	-0.0078	-0.4926	0.1328	-0.0219
ldp	0.0545	0.0301	0.0611	-0.4501	0.0713
fse	1.0000	-0.0691	0.0946	-0.0548	-0.0777
fcv	-0.0691	1.0000	-0.0124	-0.0533	0.1863
fspr	0.0946	-0.0124	1.0000	-0.2030	-0.0555
fdp	-0.0548	-0.0533	-0.2030	1.0000	-0.0148
scv	-0.0777	0.1863	-0.0555	-0.0148	1.0000
sspr	0.0931	-0.0103	0.1013	0.0105	-0.1450
sdp	0.0860	0.0008	0.0156	-0.0078	-0.0913
cspr	-0.0592	0.2363	0.0397	-0.0987	0.4719
cdp	-0.0386	-0.1430	-0.0244	-0.1655	-0.2257
spdp	0.0317	0.0552	-0.0730	-0.1697	0.0515
ldns	-0.1288	0.0959	-0.4002	0.0151	0.0140
fdns	0.1638	0.1091	0.3400	-0.1244	-0.0552
sdns	0.0960	0.0012	-0.0055	0.0005	0.1508
cdns	-0.0889	0.0389	-0.0383	-0.1337	0.3603
spdns	-0.0169	-0.0837	0.1374	-0.0213	0.0048
dpdns	-0.0004	-0.0865	-0.0073	0.0514	-0.0303

Correlation of Estimates

Variable	sspr	sdp	cspr	cdp	spdp
Intercept	0.0195	0.0230	0.0250	0.2493	0.0251
let	0.1110	0.1129	0.1091	0.1822	0.1298
let2	0.0690	-0.0444	0.0051	0.0092	0.0628
cf	-0.0451	0.0056	-0.0209	0.0453	0.1081
cf2	0.0169	-0.0073	-0.0466	0.0307	-0.0059
seq	-0.0165	-0.3304	-0.0392	0.1743	-0.0131
seq2	-0.2583	0.1176	-0.1692	0.0369	0.0206
cv	0.0371	0.1425	-0.0748	-0.4363	0.0116
cv2	0.0650	0.0798	-0.3418	-0.0859	0.0322
spr	-0.2576	0.0871	-0.2460	0.0369	-0.5071
spr2	0.1629	-0.0431	-0.1660	0.1168	-0.4138
dp	0.0621	-0.1574	0.1018	-0.3260	0.1097
dp2	0.0491	-0.1024	-0.0733	0.1315	-0.4622
dens	-0.0284	0.0136	-0.0809	-0.1690	0.0069
dens2	0.0342	-0.0319	0.0711	0.1019	-0.0452
lf	-0.0646	-0.0048	0.1335	0.1311	0.0823
lse	-0.3318	-0.4734	0.0350	0.0859	-0.0348

lcv	-0.0106	0.0881	-0.5589	-0.4459	-0.0559
lspr	-0.0626	-0.0416	-0.0401	-0.0242	-0.2682
ldp	-0.0153	0.0574	-0.0118	-0.0627	-0.0424
fse	0.0931	0.0860	-0.0592	-0.0386	0.0317
fcv	-0.0103	0.0008	0.2363	-0.1430	0.0552
fspr	0.1013	0.0156	0.0397	-0.0244	-0.0730
fdp	0.0105	-0.0078	-0.0987	-0.1655	-0.1697
scv	-0.1450	-0.0913	0.4719	-0.2257	0.0515
sspr	1.0000	-0.2995	-0.0588	0.0427	-0.0863
sdp	-0.2995	1.0000	-0.0557	-0.1787	0.0560
cspr	-0.0588	-0.0557	1.0000	-0.0470	0.1516
cdp	0.0427	-0.1787	-0.0470	1.0000	-0.0926
spdp	-0.0863	0.0560	0.1516	-0.0926	1.0000
ldns	-0.1080	-0.0428	0.0385	-0.1858	-0.0614
fdns	0.0394	-0.0096	-0.0896	-0.0070	-0.0416
sdns	0.0077	-0.0509	0.0469	-0.0778	0.0205
cdns	-0.0528	-0.1645	0.2419	0.4365	-0.0024
spdns	0.1363	-0.0480	0.0698	0.0849	-0.1790
dpdns	-0.0488	0.1495	-0.0128	0.2342	0.2711

Correlation of Estimates

Variable	ldns	fdns	sdns	cdns	spdns
Intercept	-0.0567	-0.0061	0.1229	0.2505	0.0651
let	-0.2820	0.1877	0.0313	0.0478	0.0976
let2	-0.4345	0.1863	-0.0140	-0.0773	0.1360
cf	0.1865	-0.4408	0.0072	0.1321	0.0570
cf2	-0.1530	0.2231	-0.0236	-0.0431	0.0907
seq	-0.0513	-0.0147	-0.4597	0.1022	0.0471
seq2	0.0027	0.0185	0.0967	-0.1271	-0.0282
cv	0.1351	-0.0074	0.0506	-0.7863	-0.0294
cv2	0.1060	0.0079	-0.0804	-0.1945	-0.1024
spr	0.0602	0.0510	-0.0062	-0.0307	-0.3615
spr2	-0.2466	0.0733	-0.0396	-0.0088	0.2928
dp	0.1346	0.0380	-0.0241	0.0073	0.0911
dp2	0.1366	0.0534	0.0699	-0.0619	0.1063
dens	0.0038	0.0330	-0.2984	-0.3839	-0.2565
dens2	0.0499	-0.1523	0.2077	0.2548	0.2807
lf	0.3717	-0.2293	-0.0009	0.0919	-0.1529
lse	0.2336	-0.1163	0.0137	0.0539	-0.0439
lcv	-0.0306	0.0393	0.0055	-0.1216	-0.0030
lspr	0.5631	-0.2466	0.0347	0.0282	-0.1061
ldp	-0.0962	0.0392	-0.0429	0.0530	-0.0220
fse	-0.1288	0.1638	0.0960	-0.0889	-0.0169
fcv	0.0959	0.1091	0.0012	0.0389	-0.0837
fspr	-0.4002	0.3400	-0.0055	-0.0383	0.1374
fdp	0.0151	-0.1244	0.0005	-0.1337	-0.0213
scv	0.0140	-0.0552	0.1508	0.3603	0.0048
sspr	-0.1080	0.0394	0.0077	-0.0528	0.1363
sdp	-0.0428	-0.0096	-0.0509	-0.1645	-0.0480
cspr	0.0385	-0.0896	0.0469	0.2419	0.0698
cdp	-0.1858	-0.0070	-0.0778	0.4365	0.0849
spdp	-0.0614	-0.0416	0.0205	-0.0024	-0.1790
ldns	1.0000	-0.5016	0.0771	-0.2204	-0.0977
fdns	-0.5016	1.0000	0.0913	0.0034	-0.1780
sdns	0.0771	0.0913	1.0000	0.0234	-0.1888
cdns	-0.2204	0.0034	0.0234	1.0000	0.0699
spdns	-0.0977	-0.1780	-0.1888	0.0699	1.0000
dpdns	-0.4817	-0.0977	-0.0056	0.2266	-0.4063

Correlation of Estimates

Variable	dpdns

Intercept	0.1530
let	0.0727
let2	0.1022
cf	0.0938
cf2	-0.0438
seq	0.0004
seq2	-0.0046
cv	-0.1289
cv2	-0.0647
spr	0.0908
spr2	-0.0465
dp	-0.2010
dp2	-0.3828
dens	-0.1245
dens2	-0.0308
lf	-0.0389
lse	-0.0764
lcv	-0.0668
lspr	-0.2278
ldp	0.1192
fse	-0.0004
fcv	-0.0865
fspr	-0.0073
fdp	0.0514
scv	-0.0303
sspr	-0.0488
sdp	0.1495
cspr	-0.0128
cdp	0.2342
spdp	0.2711
ldns	-0.4817
fdns	-0.0977
sdns	-0.0056
cdns	0.2266
spdns	-0.4063
dpdns	1.0000

Correlation Matrix for the Restricted Quadratic Regular Delivery Equation

Correlation of Estimates

Variable	Intercept	let	let2	cf	cf2
Intercept	1.0000	0.0292	-0.0156	-0.0871	0.0831
let	0.0292	1.0000	-0.8837	-0.5070	0.3750
let2	-0.0156	-0.8837	1.0000	0.4486	-0.4140
cf	-0.0871	-0.5070	0.4486	1.0000	-0.9105
cf2	0.0831	0.3750	-0.4140	-0.9105	1.0000
seq	-0.0076	-0.0201	0.0055	-0.0831	0.0865
seq2	0.0338	0.0152	-0.0041	0.0318	-0.0311
cv	-0.0625	-0.0326	0.0006	-0.1159	0.0909
cv2	0.0548	0.0033	0.0214	0.1215	-0.0966
spr	-0.0230	-0.2176	0.2240	-0.0227	0.0192
spr2	0.0047	0.2215	-0.3199	0.0206	-0.0465
dp	-0.5369	-0.4550	0.3977	-0.0525	0.0975
dp2	0.5477	0.3142	-0.3491	0.0829	-0.1197
dens	-0.2548	-0.1112	0.0750	-0.0714	0.0994
dens2	0.1708	0.1034	-0.0777	0.0470	-0.0652

Correlation of Estimates

Variable	seq	seq2	cv	cv2	spr
Intercept	-0.0076	0.0338	-0.0625	0.0548	-0.0230

let	-0.0201	0.0152	-0.0326	0.0033	-0.2176
let2	0.0055	-0.0041	0.0006	0.0214	0.2240
cf	-0.0831	0.0318	-0.1159	0.1215	-0.0227
cf2	0.0865	-0.0311	0.0909	-0.0966	0.0192
seq	1.0000	-0.8336	0.0339	-0.0590	-0.0247
seq2	-0.8336	1.0000	-0.0781	0.0834	-0.0037
cv	0.0339	-0.0781	1.0000	-0.8648	-0.1891
cv2	-0.0590	0.0834	-0.8648	1.0000	0.1362
spr	-0.0247	-0.0037	-0.1891	0.1362	1.0000
spr2	0.0553	-0.0222	0.1801	-0.1546	-0.8303
dp	-0.0734	0.0631	-0.0497	0.0595	-0.2045
dp2	0.0650	-0.0709	0.0856	-0.0775	0.0833
dens	-0.0065	0.0210	0.0774	-0.1035	0.0764
dens2	0.0333	-0.0328	-0.0305	0.0550	-0.0561

Correlation of Estimates

Variable	spr2	dp	dp2	dens	dens2
Intercept	0.0047	-0.5369	0.5477	-0.2548	0.1708
let	0.2215	-0.4550	0.3142	-0.1112	0.1034
let2	-0.3199	0.3977	-0.3491	0.0750	-0.0777
cf	0.0206	-0.0525	0.0829	-0.0714	0.0470
cf2	-0.0465	0.0975	-0.1197	0.0994	-0.0652
seq	0.0553	-0.0734	0.0650	-0.0065	0.0333
seq2	-0.0222	0.0631	-0.0709	0.0210	-0.0328
cv	0.1801	-0.0497	0.0856	0.0774	-0.0305
cv2	-0.1546	0.0595	-0.0775	-0.1035	0.0550
spr	-0.8303	-0.2045	0.0833	0.0764	-0.0561
spr2	1.0000	0.1206	-0.0687	-0.0133	0.0211
dp	0.1206	1.0000	-0.8932	0.0361	-0.0167
dp2	-0.0687	-0.8932	1.0000	-0.0704	0.0238
dens	-0.0133	0.0361	-0.0704	1.0000	-0.9208
dens2	0.0211	-0.0167	0.0238	-0.9208	1.0000

Correlation Matrix for the Parcel/Accountable Model

Correlation of Estimates

Variable	Intercept	pcl	pcl2	act	act2
Intercept	1.0000	-0.0616	0.0586	-0.1534	0.0286
pcl	-0.0616	1.0000	-0.4975	-0.2042	0.0820
pcl2	0.0586	-0.4975	1.0000	0.1834	-0.2117
act	-0.1534	-0.2042	0.1834	1.0000	-0.6879
act2	0.0286	0.0820	-0.2117	-0.6879	1.0000
dp	-0.6211	-0.4205	0.1232	-0.3081	0.2934
dp2	0.2867	0.3764	-0.3796	0.1589	0.1356
pact	-0.0757	-0.0679	-0.5768	-0.2066	0.2943
padp	0.1329	-0.6483	0.4871	0.2842	-0.1714
acdp	0.1703	0.2120	0.2874	-0.2533	-0.3437

Correlation of Estimates

Variable	dp	dp2	pact	padp	acdp
Intercept	-0.6211	0.2867	-0.0757	0.1329	0.1703
pcl	-0.4205	0.3764	-0.0679	-0.6483	0.2120
pcl2	0.1232	-0.3796	-0.5768	0.4871	0.2874

act	-0.3081	0.1589	-0.2066	0.2842	-0.2533
act2	0.2934	0.1356	0.2943	-0.1714	-0.3437
dp	1.0000	-0.6142	0.2484	0.0704	-0.1422
dp2	-0.6142	1.0000	0.3516	-0.4952	-0.5117
pact	0.2484	0.3516	1.0000	-0.5251	-0.5912
padp	0.0704	-0.4952	-0.5251	1.0000	0.0841
acdp	-0.1422	-0.5117	-0.5912	0.0841	1.0000

ADVO/USPS-T14-7. Please consider the translog model you present on page 56 of your testimony.

(a) Please provide the anti-log form of this model in equation form.

(b) Of the 1,545 zip-code-day observations available for your quadratic general and parcel/accountable delivery models, how many instances (observations) of zero volumes for one or more shapes did you find?

(c) Did you attempt to run translog regressions using the disaggregated volume shapes on the remaining zip-code-day observations? If so, please provide all results for model runs. If not, please explain why translog regressions were not attempted using the remaining non-zero volume observations.

ADVO/USPS-T14-7 Response:

a. $$DT = e^{\left(\beta_0 + \beta_1 V + \beta_2 V^2 + \beta_3 DP + \beta_4 DP^2 + \beta_5 DN + \beta_6 DN^2 + \beta_7 V{*}DP + \beta_8 V{*}DN + \beta_9 DP{*}DN\right)}$$

where:

DT = Delivery time
V = Ln(Volume)
DP = Ln(Delivery Points)
DN = Ln(Density)

b. 927 observations.

c. No. Doing so would have required eliminating 60 percent of the observations.

ADVO/USPS-T14-8. Suppose your translog specification using aggregate volume was run twice, once using the full data set as you did, and the second time run only on the data set containing non-zero volumes.

(a) Are you aware of any statistical tests that could test for the null hypothesis that both models are identical (collectively have the same coefficient values)? If so, would it be appropriate to test for a disaggregated model, considering all shape volumes as separate variables, using the reduced data set? Please explain fully.

(b) Is there any reason to believe that the translog model using the reduced aggregate volume data set would be considered biased or inconsistent? How would the percentage of zero-volume observations affect such circumstances, if at all? Please explain fully.

ADVO/USPS-T14-8 Response:

a. Yes. Under various assumptions about the structure of the error variance, such tests exist. Nevertheless, it would not be appropriate to do a subset test on the aggregate model to justify the use of only 40% of the data in estimating the disaggregated model. One cannot directly test the disaggregated model because it requires estimating the translog specification on all the data, containing zeros. In addition, it would be inappropriate to infer results about the disaggregated model from testing the aggregated model. The disaggregated model has far more and many different coefficients to be estimated than the aggregate model. In sum, a test on the aggregate model is not the same as a test on the disaggregated model.

b. Yes. Generally, the larger the percentage of zero-volume observations, the larger the concern arising from eliminating them from the estimation.

ADVO/USPS-T14-9. On page 55 of your testimony, referring to your translog model, you state:

Because the data were mean centered before estimation of the equation, the volume variability is just the first-order coefficient on the aggregate volume term.

(a) Please explain fully why you mean centered the data before estimating the translog model. Please explain fully when it is appropriate to perform regressions on the original (non-mean centered) data.

(b) In the above statement, do you mean that the volume variability should be based on the first order coefficient on the aggregate volume term only if the data are mean centered? If so, why? When is it appropriate to include the second order coefficient for the volume variability calculation when data are mean centered? Please explain fully.

(c) Please explain fully circumstances when only the first order coefficient and, separately, both the first and second order coefficients should be part of the marginal cost calculation when data are not mean centered.

(d) Please demonstrate the marginal cost calculation for your aggregate volume variable from your translog model with: a) only the first order coefficient included, and b) both the first and second order coefficients included.

ADVO/USPS-T14-9 Response:

a. Mean centering translog equations is standard practice both in both academic economic studies and Postal Rate Commission proceedings. Mean centering does not change the estimated variability but greatly simplifies its calculation. This point was made by the Commission in Docket No. R2000-1:[4]

> [3270] The Commission has reviewed this issue carefully, partly because mean-centering has been used by the Commission in the past. Postal Service witness

[4] See, PRC Op., Docket No. R2000-1, Vol. 1, at 176.

> Greene specifically states, at Tr .46E/22078, that mean-centering would give different coefficients but that the elasticities would be the same. UPS witness Neels states that mean-centering is a "computational convenience" and "shouldn't change the result." *Id.* at 21925. Short of calculating the elasticity with an erroneous formula, there does not appear to be support for the Periodicals Mailers' position.
>
> [3271] The Commission does not accept that there are difficulties with the use of mean-centered data. Using mean-centered data is a common practice. It involves dividing each point in the data by a constant, which happens to be the mean of the data set. If mean-centering changed the elasticity, then one would get different volume-variable costs by measuring the costs in cents instead of dollars, an obviously absurd implication.

b. Yes. If the data are not mean centered, then higher order terms enter the calculation of the variability. When the data are mean centered, one does not need to use the higher order term in calculating the variability. To clarify, consider a simple translog in one variable. Suppose that the equation is estimated both with and without mean centering:

Without mean centering:

$$\ln y = \delta_0 + \delta_1 \ln x + \delta_2 (\ln x)^2$$

With mean centering:

$$\ln y = \beta_0 + \beta_1 \ln\left(\frac{x}{\bar{x}}\right) + \beta_2 \left(\ln\left(\frac{x}{\bar{x}}\right)\right)^2$$

The variability is calculated by finding $\partial \ln y / \partial \ln x$ at the mean level of x:

Without mean centering:

$$\frac{\partial \ln y}{\partial \ln x} = \delta_1 + 2\delta_2 (\ln \bar{x})$$

With mean centering:

$$\frac{\partial \ln y}{\partial \ln x} = \beta_1 + 2\beta_2 \ln\left(\frac{\bar{x}}{\bar{x}}\right) = \beta_1.$$

c. A similar set of mathematics as was used to answer part b. governs the calculation of marginal cost.

d. The marginal cost is calculated by finding $\partial y / \partial x$. Below I demonstrate the calculation of marginal cost for the mean centered and non-mean centered translogs. This shows how one uses the first and second order terms.

Without mean centering:

$$\frac{\partial y}{\partial x} = e^{\delta_0}\left[e^{\delta_2 (\ln \bar{x})^2}\delta_1 \bar{x}^{\delta_1 - 1} + e^{\delta_1 \ln \bar{x}} e^{\delta_2 (\ln \bar{x})^2} \frac{2\delta_2}{\bar{x}} \ln(\bar{x})\right]$$

With mean centering:

$$\frac{\partial y}{\partial x} = e^{\beta_0} \frac{\beta_1}{\bar{x}}.$$

ADVO/USPS-T14-10. From a conceptual or specification view point, are you aware of any advantages to using a translog specification instead of the quadratic models you recommend to generate variability estimates? Or is the preference established only after generation of the statistical properties of particular models? Please explain fully.

ADVO/USPS-T14-10. Response:

Both the translog and the quadratic specifications are flexible functional forms and are typically used when there is no extra-sample information to help specify the functional form. Note that the translog is quadratic in logarithms of the variables as opposed to the levels of the variables. This means the translog and quadratic are not nested and direct statistical testing of the two functional forms is not available. Thus, the choice between the two functional forms is not generally made on a comparison of statistical properties. The choice may be based upon previous work for the industry or activity being modeled, extra-sample information about technology, or characteristic of the data set (as is the case here). Both functions are widely used in empirical studies as neither imposes restrictions on returns to scale prior to the estimation. A particular advantage of the quadratic is that it can be estimated when the right-hand-side variables take on zero values.

ADVO/USPS-T14-11. For all the alternative runs in Section G of your testimony,

(a) Please provide the estimated coefficients, HC standard errors and HC t-statistics. To report the data, please use the format you used in Table 18, page 56 of your testimony for your translog specification.

(b) Please provide all SAS logs for these alternative runs.

ADVO/USPS-T14-11:

a. Please see my response to OCA/USPS-T14-11. That response contains <u>all</u> of the regression output for the alternative regressions I investigated along the research path. All available material responsive to your question is contained in that response.

b. I did not retain the SAS logs from the alternative models

ADVO/USPS-T14-12. Please refer to page 47 of your testimony where you describe and report results for your alternative volume model.

(a) Explain fully why you did not use the alternative letters definition in your restricted quadratic specification.

(b) Wouldn't recognition of the DPS-cased letters marginal cost difference as confirmed by the model provide a more accurate distribution of total volume variable costs by shape and technology employed? Please explain fully.

ADVO/USPS-T14-12 Response:

a. I did. The results of using the alternative letters definition in the restricted quadratic specification is provided on page 48 of my testimony.

b. The selection of the model used to calculate volume variable costs is a decision along several dimensions. In the past, delivery models had been estimated with letters, flats, and parcels as separate variable to reflect the differential cost causing attributes of each of those shapes. However as things change in the methods of delivery, this specification may be reviewed. For example, in the current analysis, I extended this shape vector to include sequenced mail as a separate cost pools and to split large and small parcels into different cost pools. The specification of the letter cost pools for the regular delivery cost model is more difficult, as it is my understanding that in 2002, when the data were collected, a mix of delivery technologies was being used. The is more impact from delivery point sequencing today than there was in 2002 and the Postal Service was in the transition to a single case for letters and flats. Thus, the choice is not clear cut and I estimated both the traditional model and a model with separate cased letters and DPS letters. On balance, I recommended the model with the combined letters shape vector because it more closely aligns with what has been done in the past, reflects the mix of technologies in place when the data were collected, and provides intuitive results for all of the shape vectors.

ADVO/USPS-T14-13. Please refer to your fixed effects model described in pages 44 and 45 of your testimony.

(a) Please confirm that the methodology employed for your fixed effects model only required estimation of coefficients for each of the independent variables included in that model. If not, please explain.

(b) Please confirm that your fixed effects model explains differences in average delivery-times by zip code, fully through zip-code specific intercepts. If not, please explain.

(c) If so, please explain how your fixed effects model calculates the zip-code specific intercept values.

(d) Please explain in detail how the zip-code specific delivery times are used to calculate the shape-specific volume variabilities from the model.

ADVO/USPS-T14-13 Response:

a. I'm not sure I understand the question, but I can confirm that the fixed effects model does not require estimation of coefficients for independent variables not included in the model.

b. Not confirmed. The fixed effects model was estimated by sweeping the Zip Code specific means from each of the right-hand-side variables.

c. As explained in part b. above, Zip Code specific intercepts were not estimated.

d. The variabilities were derived directly from the estimated coefficients according to the following formula:

$$\lambda_{DT,V_i} = \frac{\partial DT}{\partial V_i} \frac{\overline{V}_i}{DT(\overline{V})} ,$$

where $DT(\overline{V})$ is the value for delivery time at the mean values for the volumes, delivery points, and density.

ADVO/USPS-T14-14. On page 29 of your testimony, you refer to a Box-Cox transformation "which can permit the estimation of a logarithmic function."

(a) Please confirm that you are referring to an estimation when zero volumes are present in the data base. If not, please explain fully.

(b) Please demonstrate the Box-Cox function mathematically and describe the non-linear properties of this estimator. Please explain fully how it estimates a logarithmic function when zero volumes are present.

a. Confirmed.

b. $g^{\lambda}(x) = \frac{x^{\lambda} - 1}{\lambda}$. This transformation is used in place of the translog because it has the following property:

$$\lim_{\lambda \to 0} \frac{x^{\lambda} - 1}{\lambda} = \ln x.$$

Note, that unlike the translog, the Box-Cox transformation is defined at zero values of x. Finally, to see non-linearity of the transformation in estimation, consider the simple model:

$$\ln y = \beta_0 + \beta_1 \left(\frac{x^{\lambda} - 1}{\lambda} \right)$$

This model is clearly non-linear in the parameters.

Response of Postal Service Witness Michael D. Bradley
To Interrogatories Posed by ADVO

ADVO/USPS-T14-15. Please provide copies of SAS data sets WORK.POOLR and WORK.POOLZ, referenced in the SAS program "Estimating Delivery Equations" from LR-K-81, in Excel file format.

ADVO/USPS-T14-15. Response:

Neither of the data sets cited in the question are input data sets nor are they output data sets. As their names suggest, they are working data sets within the SAS program. Thus, one does not need to input or output these data sets, in any format, to successfully replicate the program. As a result, I have not produced these data sets in Excel or any other format. If one would like to produce the data set in Excel format, I would recommend the following procedure:

Step 1. Use the SAS "Data" step and a "Put" statement to output the data set, including the desired variables, as a text file.

Step 2. Read the text file into Excel.

Step 3: Use the "Text to Columns" command within Excel to convert the file to an "xls" format.

CERTIFICATE OF SERVICE

I hereby certify that I have this date served the foregoing document in accordance with Section 12 of the Rules of Practice and Procedure.

Eric P. Koetting

475 L'Enfant Plaza West, S.W.
Washington, D.C. 20260–1137
(202) 268-2992, FAX: -5402
May 26, 2005