Postal Regulatory Commission
Submitted 3/28/2022 3:16:21 PM
Filing ID: 121262
Accepted 3/28/2022

BEFORE THE
POSTAL REGULATORY COMMISSION
WASHINGTON, DC 20268-0001

Periodic Reporting
(Proposal One)

Docket No. RM2022-3

PUBLIC REPRESENTATIVE COMMENTS

(March 28, 2022)

Philip Abraham (he/him/his)
Public Representative

Madison Lichtenstein (she/her/hers)
Greatly Assisting the Public
Representative

Contents

I. Executive Summary

In Docket No. RM2022-3, the Postal Service proposes a new methodology for calculating city carrier, letter route, street time variabilities and marginal costs using a single Top-Down Model. The cost pool for street time letter city carrier costs is $12.9 billion, so even small inaccuracies in estimating attributable costs for different products can lead to materially inaccurate costs. The analysis provided by the Postal Service in Proposal One is thorough and indicates a great deal of care and research was conducted in order to assemble, clean, analyze, and perform sensitivity checks on street time data to accurately assess variabilities and marginal times. The Public Representative believes that Proposal One should be accepted with a minor modification to eliminate the imputation procedures proposed by the Postal Service.

The primary data supporting the analysis is derived from a Delivery Data Set with volume and street hour data that combines operational data captured from various Postal Service data reporting systems and merged in a manner approved by the Commission in recent proceedings. This data is largely the product of Commission directives and precedent in previous variability analyses. The Postal Service also conducts a special Collection Study to measure collection volumes and incorporates new variables, such as density and ZIP Code land area data from the U.S. Census, in its analysis. The data underlying the Top-Down Model improves the accuracy and completeness of the current methodology, especially in terms of scope of the underlying data. Furthermore, the suggested estimation procedure using a correlated random effects model successfully addresses unobserved heterogeneity in estimating variabilities and overcomes

shortcomings identified by the Commission with fixed-effects models, which have typically been used in these analyses.

In terms of impact of the proposed methodology, overall variabilities are lower, reflecting lower variabilities (and costs) for Market Dominant products. Increases to parcel variabilities (and costs) only partially offset these decreases. The variability and attributable cost changes are significant; however, the underlying methodology is accurately specified and tested and the changes are theoretically reasonable given the large changes in the Postal Service's network.

While Proposal One could be accepted as submitted to improve the current methodology, it could be further refined. Imputation performed by the Postal Service appears to be conceptually unfounded, reduces model fit, increases resources needed to update street time variabilities in the future, and could create perverse incentives for the Postal Service to tweak some of its arbitrary assumptions for its own benefit. Also, multicollinearity does not appear to be an issue for Proposal One if no imputation procedures were conducted, and the model estimates are similar. Imputation in this analysis seems to have limited, if any benefit, incurring tangible costs, reduced transparency, and it relies on a questionable statistical basis. For these reasons, the Commission should consider accepting Proposal One without imputation as analyzed by the Public Representative in the No Imputation analysis submitted in conjunction with these comments. Lastly, while the Collection Study is at least as good as Commission-approved precedent studies and expands the scope of available collection data for modelling, there is much to be learned about collection volumes, and the Commission

should continue to direct the Postal Service to collect these volumes in special studies for future street time variability analyses.

II. Introduction

In this proceeding, the Postal Service proposes changes to analytical methods used in calculating attributable city carrier, letter route, street time costs by employing an overall top-down model of street time variability (Top-Down Model), labeled as Proposal One, filed January 5, 2022.[1] [2] During the course of this proceeding, the Postal Service provided information in response to five Chairman Information Requests (CHIRs).[3]

[1] Docket No. RM2022-3, Petition of the United States Postal Service for the Initiation of a Proceeding to Consider Proposed Changes in Analytical Principles (Proposal One), January 5, 2022 (Petition). The Petition was accompanied by a study supporting its proposal. *See* Michael D. Bradley, On the Estimation of a Top-Down Model for City Carrier Street Time, January 5, 2022 (Bradley Study). Supporting documentation was provided in public Library Reference USPS-RM2022-3-1, Public Material Relating to Proposal One, January 5, 2022. Certain files in this library reference were subsequently revised on January 21, 2022, namely "USPS-RM2022-3_1_Rev01212022.zip." The Postal Service also submitted nonpublic library references with its proposal, which the Public Representative has reviewed. *See* Library Reference USPS-RM2022-3-NP1, Nonpublic Material Relating to Proposal One, January 5, 2022 and the subsequently revised library reference Nonpublic Material Relating to Proposal One (REVISED 1-21-2022), January 21, 2022. Also *see* "Notice of Minor Revision to USPS-RM2022-3-1 and USPS-RM2022-3-NP1 – Errata," January 21, 2022.

[2] Please note that Top-Down Model refers specifically to the model presented in Proposal One. References to top-down (or single or unified) models refer to predecessors of this model or the concept of a top-down model in abstract terms.

[3] Docket No. RM2022-3, Responses of the United States Postal Service to Questions 1-3 of Chairman's Information Request No. 1, January 18, 2022 (Response to CHIR No. 1); Docket No. RM2022-3, Responses of the United States Postal Service to Questions 1-4 of Chairman's Information Request No. 2, February 15, 2022 (Response to CHIR No. 2); Response to CHIR No. 2 also included an attachment, "ChIR.2.Resp.Attachmnt.SAS Programs.zip" (CHIR No. 2 Attachment); Docket No. RM2022-3, Responses of the United States Postal Service to Questions 1-6 of Chairman's Information Request No. 3, February 22, 2022 (Response to CHIR No. 3); Response to CHIR No. 3 also included an attachment, "OneDrive_2022-02-22.zip" (CHIR No. 3 Attachment); Docket No. RM2022-3, Responses of the United States Postal Service to Questions 1-19 of Chairman's Information Request No. 4, March 11, 2022 (Response to CHIR No. 4); Response to CHIR No. 4 also included an attachment, "ChIR4.Q19.Resp.xlsx" (CHIR No. 4 Attachment); Docket No. RM2022-3, Responses of the United States Postal Service to Questions 1-11 of Chairman's Information Request No. 5, March 14, 2022 (Response to CHIR No. 5); Response to CHIR No. 5 also included an attachment, "ChIR5.Q8.MTs.Var.w.ClustrCoeffcnts.zip" (CHIR No. 5 Attachment) and was filed in association with nonpublic Library Reference USPS-RM2022-3-NP2.

Pursuant to 39 U.S. C. § 505, the Postal Regulatory Commission designated the Public Representative to represent the "interests of the general public" in the instant proceeding. According to 39 C.F.R. § 3050.11(a), the Commission should accept Proposal One if it improves the quality, accuracy, or completeness of the data or analysis of data contained in the Postal Service's annual periodic reports to the Commission. The Public Representative reviewed the submitted Petition, Bradley Study, Library References, and Responses to CHIRs and offers the following comments to aid the Commission in making a determination on the merits of Proposal One and whether its approval would be in compliance with 39 C.F.R. § 3050.11(a).

The main objective of Proposal One is to update and improve the methodology for calculating attributable city carrier, letter route, street time costs. Petition at 2. The total street time cost pool for FY 2021 was $12.9 billion. Petition at 4. In Proposal One, the Postal Service presents the results of a study of city carrier street time costs using expanded operational carrier data and estimates street time labor variabilities and marginal times for delivering certain products (referred to henceforth as marginal times) in a unified Top-Down Model. Petition at 2. The Public Representative believes that the Commission's approval of Proposal One as it is currently submitted would be in compliance with 39 C.F.R. § 3050.11(a), but the Public Representative also suggests that foregoing the complex and unnecessary imputation procedures performed in Proposal One would improve the proposal for conceptual, empirical, and practical reasons.[4]

[4] Imputation is a method of dealing with missing data in which missing values are estimated based on available information. In this case, single imputation (for specific missing values) is performed. *See e.g.,* Khan, S.I., Hoque, A.S.M.L. SICE: an improved missing data imputation technique. *J Big Data* 7, 37 (2020). https://doi.org/10.1186/s40537-020-00313-w.

III. Background and Previous Commission Dockets

a. Docket No. RM 2017-5

The Commission approved the Postal Service's proposal for estimating updated variabilities for city carrier street time in Docket No. RM2015-7, which consisted of three separate econometric equations to estimate variabilities: one for letters and flats, one for in-receptacle parcels, and one for deviation parcels and accountables.[5] [6] Petition at 2; Order No. 2792 at 66.[7] One of the major obstacles for estimating a single model was data availability, and the Postal Service needed to run special studies to collect volume measures for in-receptacle parcels, deviation parcels, accountables, and collection mail. *Id*.

For example, the Commission reviewed data from the FY 2013 Collection Mail Study, in which the Postal Service completed a special study of city carrier mail collections of 300 ZIP Codes for twelve consecutive delivery days from Monday, April 29, 2013 through Saturday, May 11, 2013. Report on the City Carrier Street Time Study, USPS-RM2015-7/1, Docket No. RM2015-7, December 11, 2014 (RM2015-7 Report) at 32; Order No. 2792 at 32. The FY 2019 Collection Mail Study had a participation rate of 98.6 percent. *Id.* In its analysis, the Postal Service used standard imputation techniques to fill in missing or incomplete data. *Id.* The Commission offered the following summary of the

[5] Unless otherwise noted, all variabilities refer to street time variabilities – how street time changes in proportion to volume changes. *See* Bradley Report at 44.

[6] For definitions of postal terms, please refer to USPS.com, About, Postal Terms, accessed at https://about.usps.com/publications/pub32/pub32_terms.htm, March 13, 2022. For example, "accountable mail" is defined as "mail that requires the signature of the addressee or addressee's agent upon receipt to provide evidence of delivery or indemnification for loss or damage."

[7] Order No. 2792, Order Approving Analytical Principles Used in Periodic Reporting (Proposal Thirteen), Docket No. RM2015-7 October 29, 2015.

RM2015-7 Public Representative's critique (in that docket) of the FY 2013 Collection Mail Study:

> With respect to the duration of the study, the Public Representative asserts that for the study to be meaningful, "the underlying data needs to span a time frame adequate for the analysis of changes in the level of cost drivers." She claims the Postal Service has offered no evidence to support a conclusion that volume variability produced by peak or trough volume can be captured by the limited data collection time frame.
>
> Order No. 2792 at 19 (citations omitted).

Ultimately, while the Commission voiced support for development of more comprehensive data for the collection study as well as delivery data, it found the sample period of twelve consecutive days acceptable because more comprehensive data did not exist. *Id.* at 59. It commented that the Postal Service's "proposed study [for Proposal Thirteen] was successful in obtaining data largely comprised of consecutive days." *Id.* While the Commission approved the imputation techniques used by the Postal Service in that docket, it critiqued imputation performed in analyses submitted by other parties. For example, it stated that "The Commission finds the need to impute volume data that are not otherwise available substantially undermines the reliability of the National Form 3999 [UPS] model." *Id.* at 61.

While the Commission approved the separated variability model, it directed "the Postal Service to collect the information needed to determine whether a single model could produce improved estimates of variability." Order No. 2792 at 65. The Commission (and the Postal Service) believed that there were potential advantages of a unified (or single or top-down, used synonymously henceforth) model, which would allow the Postal Service to estimate variabilities and marginal costs simultaneously. *Id.* The Commission

also noted that, "recording reliable daily volumes of collection mail and delivered parcel and accountable pieces would allow for more frequent updates of street time variability using much larger datasets while reducing or eliminating the need for expensive and time-consuming special data collection studies." *Id.*

b. Docket No. PI2017-1

The Commission initiated a public inquiry case Docket No. PI2017-1 to follow up on its directives that the Postal Service provide updates on its data capabilities and its progress in developing a top-down model for carrier street time. Petition at 2. On August 18, 2017, the Postal Service reported its findings.[8] The Postal Service reported three overarching issues in creating a single model for street time variability analysis: the Postal Service did not measure or have a proxy for collection volumes of mail collected from customers' receptacles, its preliminary models suffered from "serious, potentially disqualifying, multicollinearity," and low accountable mail volumes led to unreliable estimates for accountable elasticity and marginal time. *Id.* at 38-39.

The Commission commented on these issues in its interim Order No. 4869.[9] On the first point, the Commission, UPS, the PI2017-1 Public Representative, and the Postal Service agreed that collection mail is "an essential, missing variable" in a single model of street time. Order No. 4869 at 7. The Commission elaborated that obtaining collection volume was a "major data obstacle to developing and testing" a top-down model of street

[8] United States Postal Service Report on Research into the Ability of a Top-Down Model to Accurately Estimate City Carrier Street Time Variabilities, August 18, 2017 (Status Report). The Postal Service filed a library reference containing relevant data discussed in the Status Report. Library Reference USPS-PI2017-1/2, August 18, 2017.

[9] *See* Order No. 4869, Interim Order, Docket No. PI2017-1 November 2, 2018.

time. *Id.* at 11. It urged the Postal Service to evaluate different options to collect these volumes, including a CCCS method, a Ratio Method, and a $24 million 12-day sample special study each quarter to reprogram Mobile Delivery Devices (MDDs) to measure collection volumes. *Id.*

In terms of multicollinearity, the Commission ordered that the Postal Service expand the scope of the underlying ZIP Code delivery and volume data in terms of geography and sampled months throughout the year, noting that "additional data helps mitigate the inefficiency of a model that results from high multicollinearity by increasing the sample variability in the explanatory variables and, thereby, reducing the standard errors of the regression coefficients." *Id.* at 12-13.[10] In Docket No. PI2017-1, the Postal Service increased the scope of the unified model delivery data from one to two months but kept its ZIP Code sample at 300. *Id.* at 13. The Commission critiqued that 300 Zip Codes only represented about 2 percent of the 16,500 active ZIP Codes served by city carriers at the time. *Id*.[11] The Commission also suggested that estimates of variability for low-volume products, like accountables, could potentially be improved by "a large sample or oversampling [of] ZIP Codes with high accountable volumes." Order No. 4869 at 15.

[10] For an introductory discussion of multicollinearity, please see "Multicollinearity and Regression Analysis," *Journal of Physics: Conference Series*, August 2017, accessed at https://iopscience.iop.org/article/10.1088/1742-6596/949/1/012009/pdf. The paper defines that, "Multicollinearity, or near-linear dependence, is a statistical phenomenon in which two or more predictor variables in a multiple regression model are highly correlated." *Id.* at 1. Also *see* Bradley Study at 54; Status Report at 13 for the Postal Service's discussion of multicollinearity in the context of estimating variability for city carrier street time.

[11] The Postal Service claimed that the correct number of ZIP Codes with city carrier routes was 11,130, not 16,500. (300/11,130)*100 = 2.7 percent. *See* First Status Report of The United States Postal Service in Response to Order No. 4869, April 19, 2019 at 3.

The overarching directive from the Commission was for the Postal Service to construct an expanded dataset of city carrier delivery data and provide status reports to update the Commission on its progress in developing said dataset. *Id.* at 17. The required expanded dataset was described as follows:

> The dataset shall include data for each of the 12 consecutive calendar months, for 1 randomly drawn regular workweek from the expanded set of ZIP-Code-days served by regular city carriers. This expanded dataset should include data for all variables that are currently part of the Model. The Commission also suggests that it would be useful to have an alternative measure of density and asks the Postal Service to provide daily ZIP Code miles for each ZIP-Code-day, if feasible.
>
> *Id*. at 16 (footnotes omitted).

Using methodologies and merging protocols approved by the Commission in Docket No. RM2019-6, the Postal Service produced a dataset for city carrier street time that included street hours, letter and flat volumes, and parcel and accountable volumes.[12] It provided this data and updates to its methodology in five status reports in response to Order No. 4869.[13] In the First Status Report, the Postal Service described its methodology for choosing the sample in the expanded city carrier delivery dataset – the relevant pools of weeks of the month and ZIP Codes. First Status Report at 1-2. Next, it outlined four stratification criteria for the delivery data based on factors that would likely

[12] *See* Order No. 5405, Order on Analytical Principles Used in Periodic Reporting (Proposal One), Docket No. RM2019-6, January 14, 2020, at 16. Please note this proposal is a different Proposal One than the instant docket and any references to this docket will mention the associated docket number.

[13] First Status Report of The United States Postal Service in Response to Order No. 4869, April 19, 2019 (First Status Report); Second Status Report of The United States Postal Service in Response to Order No. 4869, July 31, 2019 (Second Status Report); Third Status Report of The United States Postal Service in Response to Order No. 4869, October 24, 2019 (Third Status Report); Fourth Status Report of The United States Postal Service in Response to Order No. 4869, January 22, 2020 (Fourth Status Report); Fifth Status Report of The United States Postal Service in Response to Order No. 4869, February 27, 2020 (Fifth Status Report).

materially affect street time – ZIP Codes that receive FSS processing with low accountable volume, ZIP Codes that receive FSS processing with high accountable volume, ZIP Codes that do not receive FSS processing with low accountable volume, and ZIP Codes that do not receive FSS processing with high accountable volume. *Id.* at 2. The Postal Service noted that it conducted previous research on whether a ZIP Code receiving FSS mail processing was materially correlated to its street time costs, and believed that creating a strata for accountables was responsive to the Commission's recommendation to oversample from high Accountable-volume ZIP Codes.[14] First Status Report at 2; also *see* Response to CHIR No. 3, question 4.c.ii. The Postal Service explained that it was still considering appropriate variables as alternative measures of density as requested by the Commission. *Id.* at 4. It further outlined its plans to collect delivery and volume data for the 12 months starting January of 2019 and to test its data collection processes on data it collected from October through December of 2018. *Id.* at 4-5.

In the Second Status Report, the Postal Service noted that it modified its stratification strategy to remove high (and low) accountable volume as a stratification criteria because it conflated with high (and low) FSS volume. Second Status Report at 3; also *see* Response to CHIR No. 3, question 4.a. This change reflected the Postal Service's belief that the modeling issues identified in Order No. 4869 relating to low accountable volume could not be solved by the Commission's recommendation to

[14] For evidence that costs for ZIP Codes that receive FSS processing (FSS ZIP Codes) are materially different from costs for ZIP Codes that do not receive FSS processing (non-FSS zones), *see* Report on City Carrier Street Time, USPS-RM2015/1, Docket No. RM2015-7, December 11, 2014 at 76; Report on Research Into the Ability of a Top-Down Model To Accurately Estimate City Carrier Street Time Variabilities, Docket No. PI2017-1, August 18, 2017 at 6 and 25; Response to CHIR No. 3, question 4.c.

oversample from high-accountable ZIP Codes. The Postal Service also submitted the expanded delivery data it collected from January through March of 2019. *Id.* at 1.

In the Third Status Report, Fourth Status Report, and Fifth Status Report, the Postal Service submitted the expanded delivery data it collected from April through December of 2019, and it stated it made no further methodological changes to its data collection or analytical processes. Third Status Report at 1-2; Fourth Status Report at 1-2; Fifth Status Report at 1-2.

Mail collected by carriers from customers' receptacles continued to be missing from the expanded dataset and was not measured in the Postal Service's operational data systems. Bradley Study at 3. In the instant docket, the Postal Service confirmed its finding that this data is essential in constructing the Top-Down Model, stating that, "Because carriers collect mail from customers' receptacles during the regular delivery process, the volumes of this type of collected mail have been included in previous analyses of city carrier street time, and a measure of this type of volume is *required* for estimating a unified, top-down, model." *Id.*, emphasis added. As a result, the Postal Service launched a collection volume study (Collection Study) in January and February of 2021, in which carriers used handheld scanners to estimate collection volumes during their routes. *Id.* at 4. The Public Representative further describes the delivery data and collection data submitted in Proposal One in Sections IV.a.1 and IV.a.2.

c. Network Shifts

Since variabilities were last established in Docket No. RM2015-7, volume declined across the Postal Service's city carrier delivery network. For example, from 2015 through

2020, First-Class Mail declined by 10.2 billion pieces, Marketing Mail declined by 15.9 billion pieces, and market dominant mail in total declined by 28.1 billion pieces. Petition at 6. These declines were only partially offset by an increase in volume of 3.2 billion pieces of competitive mail (mostly parcels). *Id.* Meanwhile, the size of the network grew as the number of city carrier delivery points increased by 2.4 million from 2015 through 2020. *Id.* The Postal Service suggests that these large network changes may lead to lower variabilities. *Id.* at 3.

IV. Description of the Proposed Top-Down Model

a. Data Preparation

The variability of street time (or hours) seeks to understand how changes in volumes of various products affect the amount of time that city carriers on letter routes are on the street. From there, it is possible to attribute street time costs to different products. Therefore, the dependent variable in the analysis is all city carrier, letter route street time. Bradley Study at 5. The explanatory variables should include all volumes that are collected and delivered during that time, as well as any other variables that might explain variation in street time hours across ZIP Codes. *Id*. As discussed in Section III.b, the Postal Service developed its operational data and expanded its data set of delivery volumes and street hours (Delivery Data Set) throughout the proceeding of PI2017-1. *Id.* However, collection data from customers' receptacles (Collection Data Set) were obtained in a separate special study (Collection Study). *Id.* Lastly, the Postal Service used non-Postal data to incorporate variables that explain non-volume related variation in street time hours across Zip Codes. *Id.* These individual datasets were cleaned,

merged, and filtered for potentially influential observations (PIOs) to create the final panel data set (Bradley Final Data Set) used in the variability analysis submitted in Proposal One.[15]

1. Delivery Data Set

The initial (or raw) Delivery Data Set contains a sample of 1,100 ZIP Codes with ZIP Code-day data for one random week (Monday through Saturday) of each month of 2019. *Id.* The Postal Service notes that 2020 data was available, but the data "do not reflect ongoing operational practice" due to the effect of the COVID-19 pandemic on causing deviations in "normal street time procedures." *Id.* n. 9. The Postal Service describes that, "ZIP Codes were selected from a sampling frame of all ZIP Codes reporting in DOIS [Delivery Operations Information System] during fiscal year 2018. The sample of ZIP Codes was stratified to account for differences between FSS and non-FSS ZIP Codes." *Id.* at 7 (internal citations omitted). The Postal Service cites the original operational data sources of the various variables in the initial Delivery Data Set. For each ZIP Code, "Total street hours and the letter and flat volumes delivered for each day" (DOIS volume), "[v]olumes for in-receptacle parcels, deviation parcels, and accountables for each delivery day" (PTR volume), "the number and types of routes and delivery points included in each ZIP Code," and "the number of non-customer collection points, like street letter boxes" come from the DOIS, Product Tracking and Reporting (PTR) system, Address Management System (AMS), and Collection Point Management System (CPMS), respectively. *Id.* at 7. The Public Representative notes that the construction of

[15] Please refer to Library Reference USPS-RM2022-3-NP1, January 21, 2022, Revised Folder, folder "Directory 2 Construct the Analysis Data Set" (Revised Non-Public Directory 2), SAS data file "*fullzip_panel.sas7bdat*" (Bradley Final Data Set) for the resulting finalized dataset from the Bradley Study.

the Delivery Data Set is consistent with the Postal Service's updates in the Docket No. PI2017-1. *See* First Status Report at 1-2; Second Status Report at 2.

The Postal Service cleaned the raw Delivery Data Set in Proposal One by identifying ZIP Codes with either no volume and positive street hours or no street hours and positive volumes. Of the 1,110 ZIP Codes, two were dropped due to data quality issues relating to volume reporting. Bradley Study at 8-9. For ZIP Codes that had more than eight days with positive DOIS volumes and zero street hours, the Postal Service determined that there was a data reporting problem, and it eliminated said ZIP Codes. *Id.* at 9-10. After this round of cleaning, the Delivery Data Set had 78,120 observations covering 1,085 ZIP Codes. *Id.* at 11. The Postal Service then identified three other potentially problematic sets of data. It found 9 ZIP Code-days with zero letter and flats volume, 294 ZIP Code-days with missing in-receptacle parcels and missing deviation parcels, and 112 Zip Code-days with zero street hours. *Id.* at 11-12. In total, 334 or 0.4 percent of the 78,048 ZIP Code-days were further identified as potentially problematic based on unusual volume and street hour reporting. *Id.* at 13.[16] The Postal Service categorized these problematic observations into four cases: "Zero Street Hours, Zero DOIS Volume, No PTR Volume Reported" (Case 1), "Zero Street Hours, Non-Zero DOIS Volume, No PTR Volume Reported" (Case 2), "Zero Street Hours, Non-Zero DOIS Volume, Non-Zero PTR Volume" (Case 3), and "Non-Zero Street Hours, Non-Zero DOIS Volume, Missing PTR Volumes" (Case 4), respectively. *Id.* at 14-20. 99.6 percent of ZIP Code-days appeared to have no reporting issues. *Id.*

[16] The Public Representative notes that this number is different than the 78,120 number referenced above. The Public Representative seeks clarification for the reason for this discrepancy.

Relating to Case 1, there are ZIP Code-day observations associated with 2 ZIP Codes. *Id*. All observations were kept as they were associated with extreme weather events, and so it makes sense there were days with no street hours, PTR volume, or DOIS volume. *Id.* at 14.

Relating to Case 2, there are ZIP Code-day observations associated with 44 ZIP Codes. *Id.* Observations associated with 5 ZIP Codes were kept as they were associated with extreme weather events. *Id.* The remaining 39 ZIP Codes were deleted because there was no discernible reason that there were days with positive PTR volume despite zero street hours and zero DOIS volume. *Id.* at 15.

Relating to Case 3, there are ZIP Code-day observations associated with 27 ZIP Codes. *Id.* Observations associated with 2 ZIP Codes appeared to be legitimate observations influenced by extreme weather events. *Id.* at 16. Certain observations were associated with 5 ZIP Codes that were deleted according to the Postal Service's analysis of Case 2. *Id.* For observations associated with 4 Zip Codes with 3 or more days of missing street hours, the Postal Service deleted the Zip Code. *Id.* For the remaining 13 ZIP Codes with only one day of missing street hours and 3 ZIP Codes with only two days of missing street hours, the Postal Service imputed street hours for the 19 ZIP Code-day observations [13 Zip Codes * (1 missing days / ZIP Code) + 3 Zip Codes * (2 missing days / ZIP Code)]. Relating to the decision to impute ZIP Codes, the Postal Service states that, "Review of those 4 ZIPs with many days with missing street hours suggested that they suffer from ongoing data reporting problems and they should be dropped. In contrast, the other 16 ZIP Code-days appear to be one-offs and the associated ZIP Codes almost

always report valid data." *Id.* The Public Representative discusses the imputation procedure in more detail in Section IV.a.3.

Relating to Case 4, there are ZIP Code-day observations associated with 55 different ZIP Codes, but only 33 of the ZIP Codes would remain after the ZIP Code deletions associated with the analyses of Cases 1 through 4.[17] The Postal Service states that, "there is a clear demarcation between the several ZIP Codes that have just a few days of missing PTR volume (3 or less days) and the small number of ZIP Codes that have many days of missing PTR volume (4 or greater)." Bradley Study at 21. For 8 ZIP Codes with 4 or more days of missing street hours, the Postal Service deletes the ZIP Code. *Id.* For 26 ZIP Codes with three or less days of missing street hours, the Postal Service imputes the PTR volume for 42 ZIP Code-days. *Id.* The Public Representative discusses the imputation procedure in more detail in Section IV.a.3.

The final Delivery Data Set has 75,888 ZIP Code-day observations from 1,054 ZIPs. Id. at 36. This marks a material increase (3.5x) in represented ZIP Codes and scope of data from Docket No. RM2015-7, where the Commission accepted a sample of 300 ZIP Codes.[18]

2. Collection Data Set

City letter carriers spend time collecting a "material" amount of mail from customers' receptacles. *Id.* at 24. The Postal Service states that for this reason, "it is

[17] Please refer to USPS-LR-RM2022-3-NP1, folder "Directory 2 Construct the Analysis Data Set," SAS log files, "Identify Delivery ZIPs with Data Issues.log" and "Impute Delivery Volumes.log." The 33 figure is cited as 34 in the Bradley Study at 20. The Public Representative seeks clarification of this discrepancy. .

[18] 1,054 / 300 = 3.51

important to include some measure of this collected volume to avoid omitted variables bias." *Id.* at 24-25 (citing Status Report at 7). In addition, as discussed in Section III.b, the Postal Service does not collect these volumes in its regular operational data. During the two week period of January 25th, 2021 through February 6th, 2021, the Postal Service conducted a special study (Collection Study) to collect these volumes, involving carriers from over 1,000 ZIP Codes logging their parcel collections using MDDs once they returned from the office. *Id.* at 25-26. The Postal Service states that "Collection points with barcodes from the…CMPS like blue boxes, wall units, firms, or mail chutes were not included in the study" because regular (non-Special Purpose Route) city letter carriers spend an immaterial time collecting mail at these points. *Id.* at 25, n. 21. Of the 1,100 ZIP Codes included in the dataset submitted in Docket No. PI2017 and the raw Delivery Data Set, only data for 1,086 ZIP Codes was represented in the Collection Study due to certain ZIP Codes not having access to correct MDDs. *Id.* at 26, Figure 9. This sample represents more than a threefold increase in the ZIP Codes surveyed in the FY 2013 Collection Mail Study.[19]

City carriers used three methods to estimate collection volumes. First, city carriers estimated collection volumes when they were sufficiently small (e.g. less than one inch of mail). *Id.* at 26. Second, city carriers could estimate collection volumes using the "flat tub method", in which the number of inches of mail placed in a flat tub is converted to volume using the "CMPS official flat tub conversion standards." *Id.* Third, city carriers could estimate collection volume using the "pincher method", in which the approximate

[19] 1,086/300 = 3.62

inches of mail as measured by the MDD device (as a sort of ruler) is converted to volume using the "CMPS pincher conversion standards." *Id.*[20]

The Postal Service reports that "A number of steps were taken to ensure that the resulting data set included all reliably collected volumes [for city carriers]." *Id.* at 27. The Bradley Study describes several modifications necessary to clean the raw Collection Data Set. *See id.* at 27-29. These modifications included the elimination of non-operationally feasible route day data with collection volumes of more than 400 collection pieces, 11 single-route ZIP Codes that did not have observations for all 12 days of the study, and data for 2 ZIP Codes that never reported collection volumes. *Id.* at 29-30; *see also* Response to CHIR No. 5, question 2. Then, 1,073 out of 1,086 ZIP Codes participated in this study, implying a participation rate of 98.8 percent. *Id.* at 29.

The Postal Service identified problematic ZIP Codes in the collection data, defined as "Any ZIP Code day for which the collection volume for the ZIP Code is less than the number of routes in the ZIP." *Id.* at 30. It found that many of the ZIP Codes were clustered in the Northeast region, and it believed that a severe weather event in early February 2021 called the "Ground Hog Day Nor'easter" may have impacted the ability of carriers in this region to collect mail from customers' receptacles. *Id.* The Postal Service affirms that it would be inappropriate to include these problematic observations as is because they were exogenously affected by the one-off weather event and would not be effective proxies for the corresponding delivery volumes in 2019 in the Delivery Data Set. *Id.* at 31. It decides to undergo imputation "for all ZIP Codes in the northeast part of the

[20] One inch of letters using these standards would convert to 25 pieces and 19 pieces, respectively. Bradley Study at 26-27.

count[r]y on February 1st and February 2nd for which the ZIP Code's recorded collection volume was less than its number of routes." *Id.* It also imputed data for "ZIP Codes which have only one or two inadequate or missing collection volume days." *Id.* at 34. The Postal Service notes that the final Collection Data Set contains 5,986 ZIP Code-day observations at 997 ZIP Codes. *Id.*

3. Imputation Procedures

The Bradley Study justifies imputation in order to preserve as many observations as possible. Bradley Study at 22, 36. This is potentially seen as needed due to the limited scope of the Collection Data Set which lapsed only 12 days. The Postal Service supports this by claiming imputation is both necessary and good as long as there are "sufficient volume data on the other days to form a reliable imputation." Response to CHIR No. 5, question 5, part b; *see also* Response to CHIR No. 1, questions 1, 2 and Response to CHIR No. 5, question 6. The Postal Service expands that, "As the number of imputations was small, it likely would have been possible to estimate the top-down model without employing them but, on balance, the imputations made a positive contribution to constructing the analysis data set." Response to CHIR No. 5, question 6.

The imputation process employed in the Bradley Study consists of imputing volumes in both the Collection Data Set and the Delivery Data Set. In the subsequent sections a comparison analysis is conducted in which modifications to this imputation process are employed to evaluate its necessity.

Upon cleaning the Delivery Data Set, the Postal Service imputes data in Case 3 and Case 4: "Zero Street Hours, Non-Zero DOIS Volume, Non-Zero PTR Volume" and "Non-Zero Street Hours, Non-Zero DOIS Volume, Missing PTR Volumes," respectively.

Bradley Study at 15, 20. Due to its research findings that volumes and street times vary substantially by day of the week, it chose to impute values for observations missing street time and volumes using the values from the other observations in the analysis in the same ZIP Code on the same day of the week. Bradley Study at 16-24. For example, the Postal Service notes that "Mondays are heavy [high-volume] days and Saturdays are light [low-volume] days. *Id.* at 17; *see* also Response to CHIR No. 4, question 7. Also, given the strong seasonal peak effects of volumes and street hours in December, it applied a multiplicative factor to imputations for observations missing December values. *Id.* In Case 3, "the formulas for computing the imputed street hour values [$\widehat{SH}_{ik}$] are given by:

$$If\ Month = \begin{cases} Jan. - Nov. \quad then & \widehat{SH}_{ik} = \frac{\sum_{j \neq k} SH_{ij}}{11} \\ Dec. \quad then & \widehat{SH}_{ik} = 1.083 * \left(\frac{\sum_{j \neq k} SH_{ij}}{11}\right) \end{cases}$$

where:" i = day of week, j = month, and k = month. *Id.* at 20. In Case 4, "ZIP Codes with 3 or less days of missing data" receive imputed PTR volumes for accountables (ACCT), in-receptacle parcels (IRP), and deviation parcels (DEVP). *Id.* at 21. Their respective imputation formulas follow:

$$\widehat{ACCT}_{ik} = \frac{\sum_{j \neq k} ACCT_{ij}}{11}$$

$$If\ Month = \begin{cases} Jan. - Nov. \quad then & \widehat{IRP}_{ik} = \frac{\sum_{j \neq k} IRP_{ij}}{11} \\ Dec. \quad then & \widehat{IRP}_{ik} = 1.46 * \left(\frac{\sum_{j \neq k} IRP_{ij}}{11}\right) \end{cases}$$

$$If\ Month = \begin{cases} Jan. - Nov. \quad then & \widehat{DEVP}_{ik} = \frac{\sum_{j \neq k} DEVP_{ij}}{11} \\ Dec. \quad then & \widehat{DEVP}_{ik} = 1.60 * \left(\frac{\sum_{j \neq k} DEVP_{ij}}{11}\right) \end{cases}$$

where i = day of week, j = month, and k = month. *Id.* at 20, 24.

Upon cleaning the Collection Data Set, imputation occurs in three cases. First, when one or two route-days misreported volumes and at least five of the other route-days had positive volumes, then the average of the valid volumes replaced the misreported volumes. Bradley Study at 28; *see also* Response to CHIR No. 5, question 5. This did not impute missing volumes nor zero volumes.[21] Second, the "Ground Hog Day Nor-'easter" resulted in numerous ZIP Code-days having questionable or missing reported volumes on February 1st and February 2nd of 2021. This resulted in imputation for ZIP Codes reporting 0 street hours and 0 collection volume those two days along with "ZIP Codes in the northeast part of the count[r]y on February 1st and 2nd (of 2021) for which the ZIP Code's recorded collection volume was less than its number of routes." Bradley Study at 31. The imputation method implemented consisted of replacing the misreported volumes with those from exactly a week prior, acknowledging collection volumes vary per day of week. *Id.* at 33. Lastly, other instances in which ZIP Codes had less than 3 questionable collection volume days experienced the same imputation process as those affected by the "Ground Hog Day Nor-'easter." *Id. a*t 34.

4. Data Merge

The final Delivery Data Set has 75,888 ZIP Code-day observations from 1,054 ZIPs. *Id.* at 36. The Postal Service notes that the final Collection Data Set contains 5,986 ZIP Code-day observations at 997 ZIP Codes. *Id.* The Postal Service matches by ZIP Code to merge the collection volumes into the Delivery Data Set. *Id.* at 35. Specifically, it describes that "the average value for collection volume for each of the six delivery days

[21] Please refer to Library Reference USPS-RM2022-3-NP1, January 5, 2022, folder "Directory 1 Collection Volume Study" (Non-Public Directory 1), Excel file "*Final Imputed.xlsx*."

of the week, across the two weeks, was calculated separately… This means that collection volumes vary across ZIP Codes and days of the week, but for each day of the week, for each ZIP Code, the collection volume amount will be the same across all months." *Id.* at 36. The Postal Service addresses that "Given that the collection data does not come from the full year, it is not possible to incorporate seasonal variation in this type of volume." *Id.* at 35.

The merging of the datasets was not perfect due to not all eligible ZIP Codes in the Delivery Data Set participating in the Collection Study and the eliminations described above in both the Collection Data Set and the Delivery Data Set. *See* Sections IV.a.1 and IV.a.2. However, the Postal Service found that "92.5 percent of delivery data ZIP Codes were matched, and 97.8 percent of the collection data ZIP Codes were matched." *Id*. at 37. The Postal Service notes that the combined collection and delivery dataset has 70,200 observations from 975 ZIP Codes. *Id.* at 38.

5. Non-Postal Data

The areas that carriers must cover in a particular ZIP Code will affect the amount of street time in a particular ZIP Code. *Id.* at 39. This is a source of non-volume variation which will determine street hours and should be controlled for in a single model of street time. Therefore, the Postal Service pulls land square miles data for each ZIP Code from the Census Bureau. *Id.* The Postal Service included a transformation of this variable (delivery points per square mile) in its street time study from RM 2015-7, and the Commission approved this methodology. Order No. 2792, Table II-2.

The Postal Service also introduces another variable to control for non-volume related heterogeneity in street hours: congestion. The Postal Service hypothesizes that "More congestion in a ZIP Code could require additional time to deliver the mail than in another ZIP Code, with the same amount of mail, and the same number of delivery points." *Id.* The Postal Service uses the U.S. Department of Agriculture's Economic Research Service's Rural-Urban Commuting Area (RUCA) standard as a proxy for congestion. *Id.* The RUCA standard classifies ZIP Codes on a scale from 1 through 10 from urban to rural based on calculations of population density, urbanization, and daily commuting. *Id.* After merging the non-postal data with the combined delivery and collection data, the Bradley Final Data Set has 70,056 ZIP Code-day observations from 973 ZIP Codes.[22] *Id.* at 41. This still represents more than a threefold increase compared to the 300 ZIP Code sample approved in Docket No. RM2015-7.[23]

6. Potentially Influential Observations (PIOs)

As noted in the Bradley Study, due to the large dataset (consisting of 70,056 ZIP Code-days), it is unlikely any one observation would have a significant impact on the estimates. *Id.* at 103. However, the Postal Service uses the Cook's D statistic to determine whether potentially influential observations exist and to what extent they actually affect the estimates. Specifically, ZIP Code-days with Cook's D statistics greater than 0.1 percent are considered PIOs in the Bradley Study.

[22] The Postal Service notes that two ZIP Codes must be dropped because "Two ZIP Codes in the Postal Service data set are not included in the Census Bureau file of square miles by ZIP Code." *Id.* at 41.

[23] 973/300 = 3.24

Please refer to Library Reference USPS-RM2022-3-1, January 21, 2022, Revised Folder, folder "Directory 2 Construct the Analysis Data Set" (Revised Public Directory 2), SAS output file "*Investigate PIOs with High.CooksD.lst*" (Cook's D – Bradley). 4 ZIP Code-days are concluded to have such Cook's D statistics, two of which with values relatively proximate to the 0.1 cutoff. After considering the reported street hours and volumes, the relevant weather conditions, the large sample size, and the proximity of the statistics to the cutoff, the Bradley Study concluded none of the 4 PIOs were abnormal enough to necessitate them being dropped from the dataset. *Id.* at 107, 108. It is interesting to note this was an educated conclusion, rather than one that solely relied on the statistics themselves when determining the influence of these PIOs.

b. Descriptive Analysis

1. Variables

Because the purpose of Proposal One is to estimate the "variability [and marginal times] of city carrier street time with respect to various volumes collected and delivered by those carriers," the dependent variable of the Top-Down Model is the "amount of street time incurred by all carriers in an individual ZIP Code on a given day." *Id.* at 44, 63. The main explanatory variables of the Top-Down Model are the volumes collected and delivered by city carriers: "DPS [Delivery Point Sequenced] mail, cased mail (letters and flats combined), sequenced mail, FSS mail, parcels that fit in the customers' mail receptacles, parcels that require a deviation in delivery, accountables, and mail collected from customers' receptacles." *Id.* at 44. The other explanatory variables are other non-volume factors that cause differences in daily street hours relating to ZIP Codes' individual characteristics. *Id.* In the Top-Down Model, these variables include:

> [N]umber of delivery points in a ZIP Code, a ZIP Code's square miles of land area, the proportion of business deliveries in the ZIP Code, an indicator of whether the ZIP Code is primarily a walking or driving ZIP Code, and the proportions of delivery points by type (such as door, curb, cluster box, or central deliveries).
>
> *Id.* at 46.

2. Functional Form

The Postal Service proposes a restricted quadratic functional form for the Top-Down Model. It notes that "Based upon its advantages and its track record, a quadratic functional form is appropriate for the top-down city carrier street time variability model." *Id.* at 48. The Commission approved a quadratic functional form in the last variability analysis of street time hours in Docket No. RM2015-7. Order No. 2792 at 66. The Commission noted in that docket that a quadratic functional form was a "flexible" functional form. *Id.*, Appendix B at 2.[24] The quadratic functional form was also approved by the Commission for estimating the variabilities of SPR carrier time in a more recent case.[25] Finally, there were no objections to this quadratic form being used in the Commission's interim report commenting on the Postal Service's Status Report. Order No. 4869 at 6.

The quadratic functional form is described by these parties as flexible because it "places no restrictions on the first and second order derivatives. Thus it is agnostic, *a*

[24] The Commission described this type of functional form in Order No. 2792:

> An explanatory variable in a flexible quadratic functional form regression is comprised of many terms. Specifically, it is comprised of a linear term, its square, and the cross-products of the linear terms associated with the other explanatory variables in the regression.
>
> Order No. 2792, Appendix B at 1.

[25] *See,* Order No. 5405, Order on Analytical Principles Used in Periodic Reporting (Proposal One), Docket No. RM2019-6, January 14, 2020, at 2.

priori, about the absence or presence of scale or network economies that cause the variabilities to be less than one hundred percent."[26] In other words, with this functional form the Postal Service need not make assumptions about whether there are second-order effects of changes in volume on changes in street hours; it will estimate these effects regardless and report whether they are significant.[27]

The quadratic functional form also has the benefit of not requiring observations with values of zero to be dropped, as would occur using a logarithmic or trans-logarithmic model.[28] This preserves some observations in the underlying data because certain volume and street time data in the Top-Down Model likely will plausibly (and do) take on a value of zero on certain days, especially for smaller one-route ZIP Codes and in the case of severe weather events (*See* Sections IV.a.1 and IV.a.2).

The Postal Service does not propose a full quadratic functional form, but rather a restricted functional form that "does not include cross products for the characteristic variables or between the volume and characteristic variables." Bradley Study at 49. The Postal Service notes that this restriction reduces the number of estimated coefficients in the Top-Down Model to drop from 135 to 54 and alleviates multicollinearity concerns. *Id.* The Public Representative notes that the Commission accepted a similarly restricted "Flexible Quadratic with most interaction terms." functional form in Docket No. RM2015-7. Order No. 2792, Table II-2. It also recommended estimating restricted quadratic

[26] *See* RM2015-7 Report at 25. *See* also Response to CHIR No. 3, question 1.b.

[27] The Public Representative assumes a statistical significance threshold of 0.05 as is common convention. *See* Bradley Study at 95.

[28] *See* A New Study of Special Purpose Route Carrier Costs, Docket No. RM2019-6, June 21, 2019 at 53.

models to alleviate multicollinearity concerns in the recent Docket No. RM2019-6. *See* Order No. 5405 at 27, 34; also *see* Bradley Study at 69-71.

3. Descriptive Statistics of Street Time Data

In this section the Bradley Final Data Set is sub-sampled based on Region and FSS variables to inspect how the descriptive statistics for these subsamples differ from those for the unconstrained dataset.[29] Please refer to Library Reference USPS-RM2022-3-1, January 21, 2022, Revised Folder, folder "Directory 1 Collection Volume Study" (Revised Public Directory 1), SAS output file "*Pooled Model Combined Acct Restricted Quad.RUCA Variable.Boxes.lst*." For the combined dataset, Table 1 illustrates there are 70,056 observations, the average daily street hours per ZIP Code is about 93.82, about 19.7 percent of ZIP Codes have FSS machines (FSS ZIP Codes), and on average ZIP Codes are in Metropolitan commuting areas.[30]

Next, the Postal Service conducts a summary analysis of the RUCA levels – which are grouped into Metropolitan, Micropolitan, Rural, and Small Town categories. Please refer to Library Reference Revised Public Directory 1, SAS output file "*Pooled Model Combined Acct Restricted Quad.RUCA Variable.Boxes.lst*." Mean street hours and product volumes (except for FSS volume) follow the same trend with Metropolitan having the largest values, then Micropolitan, then Small Town, and then lastly Rural with the smallest values. Bradley Study at 78. For FSS volume, the above trend holds, except Small Town and Rural both do not have FSS machines in this sample and, thus, have 0

[29] *See* Bradley Study at 61-62 for a discussion on some basic descriptive statistics relating to the Bradley Final Data Set.

[30] Refer to page 40 of the Bradley Study for a decomposition of the 10 RUCA levels.

average FSS volumes. *Id.* Considering the large differences found between the RUCA groupings, it is unsurprising that the corresponding model (run using the original dataset) includes the FSS Volume and RUCA variables which, coincidently, are statistically significant. As a result, it is informative to subsample based on these variables to determine the magnitude of their effect.

Please refer to Library Reference Revised Public Directory 1, SAS output file "*Pooled Model Combined Acct Restricted.RUCA Groupings.lst*." Compared to the unrestricted dataset Table 1 shows the Micropolitan subsample, for example, has fewer observations at 7,272, has lower average daily street hours per ZIP Code at about 53.82, and has about 2.0 percent of ZIP Codes with FSS machines. Since the original dataset on average represents Metropolitan commuting areas, it is logical for the Micropolitan subsample to have lower average street hours as these areas have less population and, thus, less mail to deliver on average which would, theoretically, take less time to deliver. Additionally, "FSS ZIP codes overall have a higher number of routes, hours, [and] mail volumes." Bradley Study at 6; Response to CHIR No. 3, question 4. Consequently, since Micropolitan areas deliver lower mail volumes than Metropolitan areas, on average, the above results have merit as this would imply there to be less FSS ZIP Codes in Micropolitan areas than in Metropolitan areas. Lastly, the corresponding Pooled Model concludes the FSS Volume variable to be statistically significant further supporting conducting an FSS subsample analysis to conclude this section. Note, due to the subsample's small size, these regression results are not reliable. Bradley Study at 79.

Please refer to Library Reference Revised Public Directory 1, SAS output file "*Pooled Model Combined Acct Restricted Quad.RUCA Variable.Boxes.FSS.lst*." Recall,

both Small Town and Rural RUCA regions have no FSS machines and, thus, are not included in the FSS subsample. Table 1 indicates this corresponding FSS subsample is about one fifth the size of the original dataset, consisting of 13,824 observations compared to the original 70,056 observations, and, thus, is not large enough to "mitigate multicollinearity." Bradley Study at 83. It also illustrates the resulting average daily street hours per ZIP Code of about 140.30 is greater than that of the original dataset while on average the ZIP Codes are in Metropolitan areas. These results are consistent with both Small Town and Rural regions not being included in the FSS subsample and, as mentioned above, "FSS ZIP codes overall have a higher number of routes, hours, [and] mail volumes." *Id.* at 6; Response to CHIR No. 3, question 4. Lastly, the corresponding Pooled Model concludes both the FSS Volume and RUCA variables are statistically significant further indicating both subsamples are informative in this analysis.

Table 1: Descriptive Statistics

Variables	Bradley Final Data Set	Micropolitan Subsample	FSS Subsample
Sample Size	70,056	7,272	13,824
Average Daily Street Hours per ZIP Code	93.82	53.82	140.30
Average Percent of ZIP Codes with FSS Machines	19.7%	2.0%	100.0%
Average Commuting Zone	Metropolitan	Micropolitan	Metropolitan

Sources: Pooled Model Combined Acct Restricted.RUCA Groupings.lst and Pooled Model Combined Acct Restricted Quad.RUCA Variable.Boxes.FSS.lst

c. Correlated Random Effects Model

1. Model Rationale

The Postal Service makes an argument that there is substantial unobserved heterogeneity in street hours – "that there are also important variations across ZIP Codes [street hours] that are not caused by volume variations, and the existence of these potentially unobserved differences could cause bias in the estimated volume coefficients." Bradley Study at 57. If these sources of heterogeneity remain unobserved, they would cause bias in the estimates of a pooled model. *Id.* at 58.[31] After conducting several investigations of trying to address unobserved heterogeneity by including more factors that could explain street hours (such as separating FSS and non-FSS Zip Codes and adding a congestion variable), it remained convinced that there was unobserved heterogeneity in its pooled model estimates. It corroborates this claim, claiming that a "pooled model nor several refined pooled models were able to provide an acceptable set of marginal times and variabilities." Bradley Study at 84; *see also Id.* at 60-84.

Theoretically, the Postal Service believes that there will be many sources of unobserved heterogeneity due to "the differences in the nature of delivery in the many ZIP Codes across the Postal Service's city carrier network." *Id.* at 85. It shows this unobserved heterogeneity by displaying variation in street times for ZIP Codes with similar volumes. *Id.* at 85, also *see id.*, Figure 11. The Postal Service explains that the primary method to deal with this heterogeneity in econometric literature is to use a fixed-

[31] The Postal Service notes that a "Pooled model estimation starts with the application of ordinary least squares to the street time variability equation…without explicit recognition of the panel structure of the data." *Id.* at 60.

effects estimation procedure.[32] [33] *Id.* at 85. However, the Postal Service noted two potential flaws of fixed-effects in variability studies as described by the Commission, that the fixed-effect estimate may be correlated with volume and that it could change with time. *Id.* at 87 (citing Opinion and Recommended Decision, Docket No. R2000-1, November 13, 2000, Appendix F at 47, 49, 71). In fact, in a recent Commission decision rejecting a proposed mail processing variability study using fixed effects, the Commission found that one of the defects of the Postal Service's proposed mail processing model was the fixed effects may be varying with time. Order No. 6096 at 35-41.

The Postal Service proposes a correlated random effects (CRE) model for the Top-Down Model in lieu of a pooled model or fixed effects model.[34] It suggests that such a model would resolve the issue of unobserved heterogeneity leading to biased coefficients and overcome Commission concerns relating to a common solution to this problem – the fixed-effects method. Bradley Study at 87; *see* also Response to CHIR No. 2, question 2.a. The CRE model treats unobserved heterogeneity in the data as random variables that are determined by the volume variables. *Id.* at 88. In this way, the Postal Service

[32] For a discussion about the use of fixed-effects as an estimation method in a regression model with unobserved (latent) variable(s), *see* Bradley Study at 88; Cheng Hsiao, Analysis of Panel Data, 27-30 (2003).

[33] In the context of mail processing, the Commission recently wrote about the Postal Service's proposed use of a fixed-effects model to estimate mail processing variabilities by assuming there were fixed factors causing idiosyncratic differences between workhours of different facilities, which was not caused by volumes of mail being processed:

> "Since there might be other factors that also affect mail processing labor variabilities, the Postal Service uses the "plant" (or "facility") level fixed-effects "to account for unobserved non-volume heterogeneity among facilities." The Postal Service includes these facility-specific fixed-effects into the model, as it has done in previous dockets, to account for "unobserved non-volume heterogeneity among facilities" and to avoid the problem of "biased and inconsistent [variability] estimates." Order No. 5405, Order on Analytical Principles Used in Periodic Reporting (Proposal Six), Docket No. RM2020-13, January 26, 2022, at 33 (citations omitted).

[34] The Bradley Study explains the theoretical underpinnings of a CRE Model, and it references academic texts relating to the model and practical applications of the model. *See* Bradley Study at 87-89; Bradley Study at 87 n.81.

argues that the CRE model resolves the first Commission concern relating to fixed-effects methods in that the CRE model assumes an explicit correlation between unobserved effects and volume. Specifically, the Postal Service notes that "The correlated random effects model captures that potential influence [of unobserved effects] by relating each ZIP Code's effects to its mean volumes." *Id.* at 89. The Postal Service notes that the characteristic variables in the in the Top-Down Model are either non time-varying (square miles, the number of boxes, and the RUCA index) or have minimal time variation (the number of delivery points, the percentage of walking routes, and the proportions of curb, cluster box, and central deliveries). *Id.* at 91. Therefore, their cross-sectional means can be used in the CRE model to "control for the presence of unobserved heterogeneity and to eliminate bias." *Id*. The Postal Service also specifies that it estimates cluster-robust standard errors to account for the possibility that errors within and between ZIP Codes are heteroscedastic. *Id.* at 89-90.

In terms of the second Commission concern with fixed-effects that the unobserved effects may vary with time, the Postal Service purports this concern is addressed by including time-period dummy variables in the CRE model for the 72 time periods in the Bradley Final Data Set. *Id.* at 93-94. It also notes that the CRE model can estimate the effects of time-varying non-volume variables in addition to time-related fixed-effects. Response to CHIR No. 2, question 2.a. The Postal Service claims that in the context of the Top-Down Model, "Controlling for time period effects is important in estimating variabilities in a top-down street time model, as omitting them can lead to biased estimates. Bradley Study at 94. In its investigation of imputation, the Postal Service showed that there were significant day-of-the-week differences in street hours and

volumes, as well as December peaks. See Section IV.a.3. The CRE model also produces a test that shows that unobserved heterogeneity exists in a pooled model for five of the characteristic variables at a 0.05 significance level and one characteristic variables at the 0.10 significance level.[35] Bradley Study at 95.

2. Model Results[36]

In order to further verify the CRE model is preferred to a naïve Pooled Model (specifically, in terms of bias), a comparison of the marginal times and variabilities between the two models is presented in Table 23 of the Bradley Study. *Id.* at 96. After providing a series of examples illustrating the plausibility of the CRE model's results, the corresponding conclusion is that the "marginal times and variabilities from the correlated random effects model are unbiased and comport with operational practice during street activities." *Id.* 99. Consequently, the CRE model is preferred for the Top-Down Model.

3. Impact

The Postal Service highlights the impact of Proposal One on variabilities and marginal costs in Section VII of the Bradley Study. *See* Bradley Study at 114-121; *see*

[35] Please refer to Revised Public Directory 2, SAS output file "*CRE Model Combined Restricted Quad With Time Effects.lst*" (CRE Model – Bradley). Please also refer to variable descriptions provided in Response to CHIR No. 1, question 3.a. The cross-sectional mean coefficients for the RUCA indicator, delivery points, percent CBU delivery points, percent curb delivery points, and percent central delivery points characteristic variables are significant at a 0.05 level. The cross-sectional mean coefficient of the Ratio of Business Delivery Points is significant at a 0.10 level with a corresponding p-value of 0.0862. The Bradley Study states that there is one additional characteristic variable whose cross sectional mean is "significant at the 0.056" level. Bradley Study at 95. The Public Representative seeks clarification. The Postal Service may be mistakenly referring to the cross-sectional mean of the CPMS Collection Points ("boxes") variable, whose coefficient has a p-value of 0.5618. However, if that is the case, the Postal Service's interpretation that this variable is contributing in explaining non-volume heterogeneity between ZIP Codes would be incorrect.

[36] Please refer to the Bradley Study at 95-100 for its interpretation of the results of its model. Please refer to Response to CHIR No. 1, question 3.c., Response to CHIR No. 2, question 1, Response to CHIR No. 5, questions 3, 8 for the Postal Service's methodology for calculating marginal times and variabilities.

also Response to CHIR No. 3, question 3. The Postal Service notes that variabilities estimated in Proposal One and those estimated in Docket No. RM2015-7 cannot be directly compared due to differences in the underlying cost pools and the previously estimated variabilities also having indirect costs in street support. *Id.* at 114-116. The Postal Service summarizes that "Applying the top-down variabilities reduces the FY 2021 volume variable street time costs for letter and flat delivery, reduces the FY 2021 cost of collecting mail from customers' receptacles, and increases the FY 2021 cost of delivering parcels." *Id.* at 119; *Id.*, Table 33. The Postal Service rationalizes these variability changes by attributing them to the significant shifts occurring in the Postal Service's delivery network, with decreased letter and flats volumes and increased parcels and delivery points. *Id.* at 119-120. The unit costs correspondingly increase for parcel products and decrease for market-dominant products. *Id.* at 120; *Id.*, Table 34. As a result of Proposal One, the Public Representative notes that total volume variable and product-specific attributable costs for office and street time (not including piggyback) costs decreased by $1.612 billion from $8.552 billion to $6.940 billion.[37] This reduction was offset by a $377 million increase in attributable costs for Competitive products. *Id.* Factoring in piggybacks, attributable costs decreased by $2.178 billion.

d. Sensitivity Analyses Conducted by the Postal Service

The Postal Service performs several analyses to verify the robustness of its Top-Down Model results. It performs a two-way correlated random effects model, which corroborates the findings of its Top-Down Model and confirms the importance of

[37] *See* Library Reference USPS-RM2022-3-1, January 5, 2022, folder "Directory 4 Public Impact Workbooks," Excel file "FY21Public.ProposalOne.CostImpact.xlsx," tab "Office and Street Cost Impact."

controlling for time effects. *Id.* at 102. It also estimates a common correlated effects pooled (CCEP) model. It states that "the pooled top-down model can be re-estimated as a CCE pooled model, including the cross-sectional averages [and time-specific effects], to produce unbiased estimates of the variabilities." *Id.* at 110, 112. The Postal Service maintains that the results of the CCEP model in terms of variabilities and marginal times corroborate the robustness of its CRE Top-Down Model estimates. *Id.* at 113.

V. Public Representative Analysis

The Public Representative believes that the Commission could accept Proposal One as submitted. Many elements of Proposal One are the product of Commission-directed research into the feasibility of a unified variability model. New elements that are introduced, such as a new density variable and a new correlated random effects procedure, are the results of the Postal Service's research and the latest econometric techniques, and these additions are meaningful to deal with unobserved heterogeneity in a model of street time variability. These considerations are especially relevant given the limitations of fixed-effects methods identified by the Commission in similar variability analyses, such as the Commission's rejection of Proposal Six in Docket No. RM2020-13, in dealing with said unobserved heterogeneity. Proposal One also appears to consider and address all of the Commission's concerns related to its development of a unified model in Order No. 4869 in Docket No. PI2017-1.

While the Top-Down Model of Proposal One improves the completeness and accuracy of estimating variabilities compared to the current methodology, the Collection Data Set has serious limitations, and the Commission should continue to direct the Postal Service to improve and expand its Collection Study methodology. Lastly, the imputation

procedures advocated to preserve data in the Collection Data Set and Delivery Data Set appear to be conceptually unsound and both empirically and practically unnecessary. The Public Representative recommends that the Commission considers approving Proposal One with a methodological modification to eliminate all observations with incomplete data instead of undergoing complex and unnecessary imputation procedures.

a. Proposal One's Raw Data are the Result of Commission Directives and Previous Commission Decisions

The Petition and the corresponding Bradley Study represent the culmination of a great deal of past economic research; many elements of the proposed Top-Down Model are the direct result of the Postal Service following up on Commission directives and following the precedent of previous Commission decisions. The Postal Service kept the Commission updated on its progress of estimating a single Top-Down Model for street time as directed in Docket No. PI2017-1. *See* Section III.b. The raw Delivery Data Set is largely the product of specific Commission directives from that docket:

> Accordingly, the Commission directs the Postal Service to provide an expanded dataset of city carrier delivery data. The dataset shall include data for each of the 12 consecutive calendar months, for 1 randomly drawn regular workweek from the expanded set of ZIP-Code-days served by regular city carriers. This expanded dataset should include data for all variables that are currently part of the Model. The Commission also suggests that it would be useful to have an alternative measure of density and asks the Postal Service to provide daily ZIP Code miles for each ZIP-Code-day, if feasible.
>
> Order No. 4869 at 16 (citations omitted).

For this reason, the Public Representative agrees with the format of the raw Delivery Data Set presented in Proposal One. As the Postal Service notes, these data are from operational sources, and they have been used in other rulemaking dockets that

have been approved by the Commission, such as Docket No. RM2019-6. While there is more recent data from 2020, the data were distorted by the effects of the COVID-19 pandemic and would not reflect the ongoing operational practices of city carriers, thus, making it not a good basis for estimating variabilities. Bradley Study at 5 n.9. Nevertheless, using the 2019 delivery data in the Top-Down Model would improve the accuracy of street time variabilities because this data is more recent than the data underlying the variabilities associated with the current methodology. Given that significant shifts in the delivery network have occurred since the previous street time variability study, more recent data will be more accurate in reflecting current street time variabilities than using older data. *See* Section III.c.

Proposal One also improves upon previous variability studies by greatly expanding the number of observations available for estimation.[38] The Public Representative also agrees with the Postal Service's decision to include an additional characteristic variable, the RUCA standard index, in order to respond to the Commission's directive to provide an alternative measure of density. This variable is statistically significant in the Top-Down Model at the 0.05 level and, thus, can be interpreted as statistically significant in explaining unobserved non-volume related heterogeneity in street time and improving the completeness and accuracy of the variability analysis through its inclusion.[39]

[38] Proposal Thirteen's Regular Delivery Model included 3,485 ZIP Code-day observations. Order No. 2792, Table II-2. Proposal Thirteen's Deviation Parcel/ Accountables Models included 3,061 ZIP Code-day observations. Order No. 2792, Table II-4. The Bradley Final Data Set has 70,056 ZIP Code-day observations.

[39] Please refer to CRE Model – Bradley.

The raw Collection Data Set is similarly the product of Commission directives. The Commission and Postal Service agreed in Docket Nos. RM2015-7 and PI2017-1 that collection volume is a material and necessary element of estimating a single unified street time variability model and without a measure of such volumes, estimates of street time variability would be biased. *See* Sections III.a. and III.b. In this docket, the Postal Service estimates the extent of that bias; it finds that the variability estimates and marginal times are higher in a model where collection volumes are not used as an explanatory variable in the Top-Down Model. Response to CHIR No. 2, question 4. For example, the DPS variability increases from 7.7% to 7.8% when eliminating collection volumes as an explanatory variable. *Id.*[40]

Because the omitted variables bias discussed above presented a risk for street time variabilities calculated without collection volumes, the Commission ordered the Postal Service to investigate the possibilities of either modifying its operating systems to record collection volumes or to:

> [L]everage the widespread use of next generation hand-held scanners by carriers to record the delivery of accountable mail and produce reliable volumes by ZIP Code. Reporting volumes of collection mail by ZIP Code may be more difficult to automate, but it may be possible to produce a manual estimate based on a length or weight conversion factor.
>
> Order No. 2792 at 65.

[40] As an aside, the Postal Service uses this opportunity to assert that "the small size of the bias suggests it may be possible to update the street time variability equation in the future without having to field a special collection volume study." Id. *See* Section V.d for further discussion of this issue.

The raw Collection Data Set is the result of this investigation, and it uses length based conversion factors to estimate collection volumes.[41] These data represent the most recently available collection data and were not affected by the COVID-19 pandemic because collection does not require customer interaction and did not require altered operations. Response to CHIR No. 3, question 5.b.i. Furthermore, the sampled ZIP Codes greatly increased in the Collection Study compared to the FY 2013 Collection Mail Study (1,086 ZIP Codes versus 300 ZIP Codes). Theoretically, these data should then serve as fine, if slightly imperfect, proxies for collection volumes in FY 2019. Moreover, they should be more accurate and complete proxies for FY 2019 collection volumes than collection volumes measured in the FY 2013 Collection Mail Study. It is not disqualifying that the Collection Study has a sample period of twelve consecutive delivery days because the Commission previously accepted the Collection Mail Study with the same sample period in Docket No. RM2015-7. The Commission stated that:

> Because more comprehensive data are not available, the Commission finds Proposal Thirteen's sample period acceptable. The duration of the sample used in Proposal Thirteen is the same as in the study underlying the current methodology, so maintaining the same duration does not impact the Commission's assessment under 39 C.F.R. § 3050.11(a).
>
> Order No. 2792 at 59.

[41] The use of length-based conversion factors (the flat tub method and the pincher method) to estimate collection volumes may result in measurement error in the explanatory variable. The Postal Service confirms this. Response to CHIR No. 3, question 6.a.i. However, the Postal Service provides compelling reasons that this measurement error is likely not material – 1)These methods are similar to the ones accepted by the Commission in its review of the FY 2013 Collection Mail Study in Docket No. RM2015-7; and 2) it is unlikely that this measurement error would have a material impact on the other volume coefficients. *Id.* Additionally, the Commission recognized that collection volumes are difficult to measure and it suggested using conversion standards to measure these volumes in Order No. 2792. Order No. 2792 at 65.

Similar logic should apply when evaluating the merits of the Collection Study and the resulting raw Collection Data Set in this case. It would be ideal if daily collection volumes in 2019 were available, but they are not. The Collection Study is essentially an update (and improvement) of the special collection study already approved by the Commission in Docket No. RM2015-7. It has a similar methodology using newer technology, more than three times the number of ZIP Codes, the same sample period duration, and a slightly higher participation rate to the FY 2013 Collection Mail Study. It also represents an improvement in accuracy to the FY 2013 Collection Mail Study because it uses more recent data, which is especially important given the described shifts in the Postal Service's delivery network.

The data provided in the raw versions of the Collection Data Set and the Delivery Data Set provide a solid basis for the Top-Down Model in Proposal One. They follow the precedence and directives of the Commission's ruling in Docket No. RM2015-7 to create a top-down model of street time variability – developing the necessary delivery data from operational sources, exploring options for adding a new density variable, and following the Commission-approved methodology of updating (and improving) the FY 2013 Collection Study to record collection volumes. In the next section, the Public Representative evaluates that Proposal One overcomes the three shortcomings of a top-down model considered by the Commission in its interim report in Docket No. PI2017-1, Order No. 4869: no collection volumes, high multicollinearity, and low accountables volumes.

b. Proposal One Addresses the Shortcomings Identified in Order No. 4869

1. The Collection Study Provides Essential Data for the Top-Down Model and Improves the Quality of Available Collection Data

All parties agreed in Docket No. PI2017-1 that collection mail volume was an "essential, missing" variable in a proposed unified model of street time variability. The Commission recommended the Postal Service evaluate various methods to gather collection volumes. See Section III.b. One of the methods the Commission suggested for collecting this data was a quarterly, $24 million, 12-day sample special study to reprogram the MDDs. The Collection Study represents the Postal Service's follow up to the Commission's suggestion. Although the Postal Service's 12-day sample Collection Study was not conducted quarterly, it still represents an improvement and update to available collection volumes data, as last measured by the Postal Service and approved by the Commission in Docket No. RM2015-7. *See* Section V.a. However, the Public Representative questions why the Postal Service did not choose to conduct this study quarterly as prescribed by the Commission, and it would be helpful if the Postal Service provided the actual cost of running the Collection Study.

2. Low Risk of Multicollinearity

The Commission recommended the Postal Service to increase the available amount of analysis data to alleviate multicollinearity concerns in a unified variability model. In relation to the current methodology, Proposal One would be an improvement because "[r]ather than being limited to a small sample of letter routes from a single month as was used in the previous study, the new study includes data from a large number of ZIP Codes across all months of the year." Petition at 3. In the Status Report, the Postal Service followed the Commission's recommendation in Order No. 2792 to expand the

scope of the analysis data set in terms of months and ZIP Codes. The Bradley Final Data Set of the Top-Down Model represents more than a threefold increase in the number of ZIP Codes sampled in some of the methodologies submitted in Docket Nos. RM2015-7 and PI2017-1. Order No. 4869 at 13; Order No. 2792, Table II-2. In addition to increasing the geographic scope of ZIP Codes, the Postal Service's Top-Down Model draws on data from all twelve months of calendar year 2019, a six-fold increase from the two month methodology submitted by the Postal Service in Docket No. PI2017-1. Order No. 4869 at 13.

The result of the expanded dataset is higher statistical power. As a result, multicollinearity does not appear to pose a strong risk to this model because all of the key explanatory variables are statistically significant in the Top-Down Model.[42] Importantly, the coefficients of deviation parcel/ accountables volume variables were both statistically significant.[43] In its recent ruling of Order No. 6096, the Commission stated its view that if variables are individually and jointly significant as expected in a variability model, multicollinearity should not be too much of a concern. Order No. 6096 at 32. Given that all the primary explanatory variables in the Top-Down Model are significant, if not at a 0.05 level, then at a 0.10 level, the Public Representative believes multicollinearity should not be a grave concern. The Postal Service acknowledges that the condition index, which measures multicollinearity, for the quadratic pooled model (run as sensitivity check to the

[42] Please refer to CRE Model – Bradley. The first order coefficients of the volume explanatory variables are all significant at a 0.05 level.

[43] Please refer to CRE Model – Bradley. The deviation parcel/ accountables variable ("devpa") and its square term coefficients are statistically significant at the 0.05 level.

CRE model) is within range of the condition indexes that the Commission cited in its order in Docket No. RM2019-6.[44]

The Public Representative agrees with the Commission's assessment that multicollinearity is essentially a data problem. Proposal One should be accepted because it meaningfully alleviates the problem of multicollinearity by greatly increasing the pool of analysis data, as well as taking meaningful steps to reduce multicollinearity, such as restricting the interaction terms in its model.

Despite the Public Representative's arguments, above, suggesting a limited risk of multicollinearity, the Postal Service states that multicollinearity is a serious issue in their Top-Down Model. Bradley Study at 51; *see also* Response to CHIR No. 5, question 6. The Public Representative notes that there are other ways to reduce the impacts of multicollinearity, besides imputation. For example, the Postal Service states that it estimates cluster-robust standard errors because the regression errors within individual ZIP Codes may be correlated with one another. Bradley Study at 90.[45] It may be possible to increase statistical power and combat the effects of multicollinearity by not estimating clustered standard errors, if errors within individual ZIP Codes are not correlated with one another. Multicollinearity could also be reduced by gathering more collection data and

[44] *See* Bradley Study at 71 (citing Order No. 5405, Appendix at 3). The Public Representative encourages the Commission to consider requesting the Postal Service to calculate the condition index for the CRE model.

[45] The Postal Service also assumes that "As in the pooled model, the error structure for the correlated random effects model is likely to be heteroscedastic." This assumption should be confirmed to see if it is necessary to use robust standard errors in this variability analysis, but the Public Representative agrees that the Postal Service's assumption is likely correct.

incorporating them into future variability analyses and further expanding the sample of ZIP Codes in the analysis.

3. Low Accountables Volumes Still Require the Combined Estimation of Accountables and Deviation Parcels Variabilities

The Postal Service finds that vastly increasing the scope of the data did not alleviate the issues of estimating accountables variabilities. The Postal Service's investigations of potential models often estimate nonsensical, negative value of variability for accountables. Bradley Study at 32. The Postal Service states that "the real culprit [of the negative variability] is the exceeding[ly] small volumes for accountables relative to the other volumes in the equation. … [T]his condition makes it extremely difficult to estimate a separate variability for accountable mail in a broad top-down style model." *Id.* The Postal made similar arguments in its Status Report. *See* Status Report at 15.

The Commission in Docket No. PI2017-1 remained open to the idea that a unique variable for accountables may not be justified based on the Postal Service's future investigations, in this case in the instant docket. Order No. 4869 at 14. Despite following up on the Commission's recommendations to reduce multicollinearity by increasing the scope of the street time data and attempting to oversample ZIP Codes with high accountables volumes (*See* Section III.b), both methods appear to fail in producing operationally plausible accountables variabilities. The Public Representative agrees with the Postal Service that the "commonality in the nature of delivery for deviation parcels and accountables has been long recognized in studies of city carrier delivery." Bradley Study at 66. The Postal Service notes, and the Public Representative confirms, that these variabilities were jointly estimated in both of the Commission-approved street time

variability dockets, Dockets No. R2005-1 and RM2015-7. *Id.* citing Order No. 2792 at 5, 46-47, 66. While a special study was required in Docket No. RM2015-7 for recording delivery times and volumes for deviation parcels/ accountables, these data are now recorded separately in the PTR, and their use has been approved by the Commission recently in Docket No. RM2019-6. As such, inability of the Top-Down Model to separately estimate variabilities for accountables and deviation parcels cannot be seen as a defect, and the estimated deviation parcels/ accountables variability represents an improvement over the current methodology in terms of the underlying data quality, scope of sampled ZIP Codes, and timely relevance given recent network shifts.

In terms of providing accurate collection volumes, dealing with manageable multicollinearity, and resolving the question of whether accountables' variability can be separately estimated from that of deviation parcels, Proposal One addresses all the issues identified in Docket No. PI2017-1, and it represents at least an improvement over the methodologies approved in Docket No. RM2015-7, if not just for the fact that more recent data is analyzed.

c. Imputation Appears to Be At Least Partially Unwarranted and Unnecessary

In Docket No. RM2015-7, the Commission noted the following of UPS imputations of three volume variables:

> The parameters developed in the imputation are estimates, and each one has a range of values which can be forecast for every observation and still be considered accurate. However, UPS treats the values of the three imputed variables as exogenous or fixed in repeated samples, even though the values of these variables are estimates, and therefore not fixed in repeated samples. By not accounting for the error associated with each estimated forecast parameter,

> the National Form 3999 model introduces a high level of measurement error in its explanatory variables.
>
> Order No. 2792 at 60.

The Commission showed a preference for not approving unnecessary and over-extensive imputation procedures. The Public Representative finds that the imputation performed by the Postal Service is not conceptually sound, especially the imputation procedure for the Collection Data Set, and it asks the Commission to approve Proposal One without these imputation procedures for theoretical, empirical, and practical reasons.

1. Theoretical Concerns

The Postal Service performs different imputation procedures to clean the raw Delivery Data Set and Collection Data Set. *See* Section IV.a.3. The Postal Service justifies that these imputations require imputing values for relatively few days, but they preserve many observations. *See* Response to CHIR No. 1, question 1.b.; Response to CHIR No. 5, question 6. The Postal Service's basis for imputing in certain situations was "where there were sufficient data to support imputation and the number of zero volume days was reasonable." Response to CHIR No. 5, question 6.b. In terms of the imputation in the Collection Data Set, the Postal Service only ever had one other data point to impute volume. It is conceptually unsound that one point of Collection data for the week before or after a certain date would be an accurate proxy for the collection volume of that date or even of the true mean of collection volumes for the Zip Code. The Public Representative cannot quantify the uncertainty around the Collection Data Set imputation estimates, but it is substantial and greater than the uncertainty presented in the Delivery Data Set imputation estimates below due to the limited scope of the Collection Data Set.

Collection volume imputation seems especially unreasonable considering the lack of information the Postal Service has relating to trends and seasonality in collection volumes. *See* Section V.d.

When imputation occurs in the Delivery Data Set, the Postal Service theoretically only has 10-11 (Case 3) or 9-11 (Case 4) data points for street time and PTR volumes, respectively, from which they can derive their imputation estimates. See Section III.a.3. While a sample of 11 data points would be more accurate in estimating the true mean value of street time for PTR volume for a particular ZIP Code and day of the week, there would still be significant uncertainty surrounding such an estimate.[46] Moreover, in this docket the Postal Service estimates the street hours and PTR volumes for a specific ZIP Code-day, so the uncertainty in the Postal Service's imputation estimates could be quantified by the sample standard deviation of the other data points for the same day of the week. The Public Representative calculates the standard deviation for all nineteen of the imputed ZIP Code-day observations in Case 3 of its Delivery Data Set imputations.[47] The Public Representative calculates that it would be reasonable to expect the imputations performed by the Postal Service in Case 3 to differ from the true street hours on the imputed day by +/- 15.5 percent. The Public Representative notes that the uncertainty in these imputations reaches up to +/- 44.2 percent as the variance of the

[46] That uncertainty of the estimate of the true daily mean would on average vary (at the minimum) from the Postal Service's imputation estimate by the standard error of approximately SD_z / sqrt(11). SD_z = Sample Standard Deviation of the 11 (street hour or PTR volume) values for the imputed ZIP Code. One then divides this by the square root (sqrt) of the number of ZIP Code-days observed (in most cases, this is 11 as most ZIP Codes are only missing one day of observations) to get the expected deviation from the true population mean. *See* Investopedia, "Standard Error of the Mean vs. Standard Deviation: The Difference," accessed at: https://www.investopedia.com/ask/answers/042415/what-difference-between-standard-error-means-and-standard-deviation.asp#toc-sem-vs-sd.

[47] *See* PR-LR-RM2022-3-NP1, folder "Delivery Data Imputation Uncertainty Analysis," Excel file "PR_case3_work.xlsx." Also *see* Bradley Study at 16.

underlying 11 data points increases.[48] The Public Representative believes it is conceptually unjustified to impute values for these data because of the little amount of data available for imputation and the variance of the underlying data. While in the Delivery Data Set this uncertainty can be quantified, the uncertainty in the Collection Data Set cannot even be quantified and assessed by the Commission.

The analysis conducted by the Public Representative suggests that imputation procedures proposed by the Postal Service in Proposal One are not founded on large enough samples to support imputation, particularly for the Collection Data Set where imputation is based on only one other value. In the next section, the Public Representative shows the results of an analysis where these imputation procedures are not conducted and will argue that imputation is also not empirically justified.

2. Empirical Concerns

i. Imputation is Unnecessary in the Collection Data Set

Despite the Postal Service's belief that imputation in the Collection Data Set is necessary in order to preserve as many observations as possible, primarily consisting of replacing ZIP Code-day collection volumes with those from the same day of the prior week, the Public Representative believes that there is insufficient data for imputation estimates to be justified. Bradley Study at 22, 36. To test this hypothesis, the Public Representative modifies the imputation process discussed in Section IV.a.3. to only include imputation on the Delivery Data Set and no imputation on the Collection Data Set

[48] Please note that in Case 3 even those 3 ZIP Codes where two observations were missing street hour data, the missing values occurred on different days of the week. Therefore, all 19 missing observations had 11 data points from which imputation was performed.

and estimates street time variabilities and Marginal Costs using the CRE model. The Public Representative will refer to this modified methodology as the Partial Imputation (PI) analysis.[49] Any observations that experienced imputation in the Collection Data Set in the Bradley Study are removed from the Collection Data Set instead of being imputed. As a result, the Bradley Final Data Set contains 70,056 ZIP Code-days and the PI Final Data Set contains 64,440 ZIP Code-days. *See* Bradley Final Data Set; PI Final Data Set. This modification is justified as the Collection Data Set does not contain enough data points for imputation to be feasible as discussed in Section V.c.1.

The PI Final Data Set's summary statistics are not materially different from those of the Bradley Final Data Set. Mean daily street hours per ZIP Code is about 93.82 hours and average percent of ZIP Code-days with FSS machines is about 19.7 percent in the Bradley Final Data Set, while mean daily street hours per ZIP Code is about 96.38 hours and average percent of ZIP Code-days with FSS machines is about 19.9 percent in the PI Final Data Set. *See* Table 2. Given their respective standard deviations, these averages are very similar. Additionally, Table 2 provides summary statistics for the corresponding CRE models relating to the Bradley Final Data Set and the PI Final Data Set. Specifically, it indicates R^2 is 0.5137 and root mean squared error (RMSE) is about 671,781.08 when the CRE model is run on the Bradley Final Data Set, while R^2 is 0.5232

[49] The attached Library Reference PR-LR-RM2022-3-NP1, (PI Non-Public Analysis) folder "Partial Imputation," Directory 1 (PI Non-Public Directory 1) and Directory 2 (PI Non-Public Directory 2) contain non-public programs which remove the imputation procedure from the Collection Data Set preparation. The attached Library Reference PR-LR-RM2022-3-1, (PI Public Analysis) folder "Partial Imputation," Directory 2 (PI Public Directory 2) contains public programs which estimate variabilities in otherwise the same fashion as Proposal One. The resulting variabilities and marginal times from the PI analysis can be found at PI Public Analysis, SAS output file "*CRE Model Combined Restricted Quad with Time Effects_PI.htm*" (CRE Model – PI). The resulting finalized dataset can be found at PI Non-Public Analysis, Directory 2 (Non-Public Directory 2), folder "Data," SAS data file "*fullzip_panel_PI.sas7bdat*" (PI Final Data Set).

and RMSE is about 677,839.57 when the CRE model is run on the PI Final Data Set.[50] [51] *Id.* The increase in R^2 by about 1.8 percent illustrates that partial imputation, relative to Proposal One, improves the model's ability to explain the variation in street hours. The increase in RMSE by about 0.9 percent indicates the model's predictive ability reduces slightly. Note, the sample size decreased by about 8.0 percent when the modifications were applied, which could explain why the RMSE increased. Because RMSE increases with a decrease in sample size, all else equal, the increase in RMSE may solely be a function of the loss of additional sample size. *Id.* This inhibits any conclusions from being made based on the increase in RMSE, especially in the case of imputation where changes in sample size can represent artificial data.

Table 2: Descriptive Statistics

Statistic	Bradley Final Data Set	PI Final Data Set	NI Final Data Set
Sample Size	70,056	64,440	62,640
Average Daily Street Hours per ZIP Code (standard deviation)	93.82 (75.56)	96.38 (75.65)	98.01 (75.56)
Average Percent of ZIP Days with FSS	19.7% (39.8%)	19.9% (39.9%)	20.5% (40.3%)

[50] R^2 (R-squared) is a statistical measure that explains the amount of variation in the dependent variable (street hours) that can be explained by the explanatory variables. For example, the Commission has written that, "High or low R-squared values do not necessarily imply "good" or "bad" econometric models. A high R-squared value implies that most of the variation in workhours can be predicted by the model." Order No. 6096 at 40 (citations omitted; citing Stock and Watson, *Introduction to Econometrics*, 2007 at 126).

[51] "The RMSE statistic provides information about the short-term performance of a model by allowing a term-by-term comparison of the actual difference between the estimated and the measured value. The smaller the value, the better the model's performance." H.D. Kambezidis, *Comprehensive Renewable Energy*, 2012, accessed at https://www.sciencedirect.com/topics/earth-and-planetary-sciences/root-mean-square-error.

Machines (standard deviation)			
R^2	0.5137	0.5232	0.5245
RMSE	671,781.08	677,839.57	675,423.06

Sources: CRE Model Combined Restricted Quad With Time Effects.lst, CRE Model Combined Restricted Quad With Time Effects_NCVI – Modified.htm, and CRE Model Combined Restricted Quad With Time Effects_NI – Modified.htm

The effective results of the PI Analysis do not vary by much when compared to the Bradley Study's results.[52] For the most part, the model variables maintain their significance, or lack thereof, assuming a 0.05 significance level, in the PI analysis. *Id*. Also, as can be seen in Table 3, the CRE models for the Bradley Final Data Set and the PI Final Data Set result in very similar marginal times and variabilities. For example, the percent changes in marginal times are less than 5 percent as a result of using the PI procedure; the absolute change in marginal times is always less than two seconds. Accordingly, the variabilities are essentially identical when removing imputation in the Collection Data Set.

The initial justification for imputing the Collection Data Set was poor, considering the small amount of data on which to base imputation estimates. In terms of model fit, the increase in R^2 seems to indicate that removing imputation improves model fit. In terms of impact, Collection Data Set imputation does not have a large effect on the estimation of marginal times and variabilities. Therefore, for theoretical and empirical

[52] Please refer to CRE Model – Bradley. Please also refer to CRE Model – PI. The only perceivable differences include dum13 (representing 03/23/2019) having a visibly different magnitude (although, it is not statistically significant, so it is not believed to have any meaningful significance) and the intercept no longer being statistically significant at the 0.0001 level (with a p-value equal to 0.0001). The Bradley Study did not mention which significance level was used as a cutoff point in its analysis, and since this p-value is well below 0.1, 0.05, and 0.01, which are the most common cutoff points, it is believed the intercept is still considered statistically significant.

reasons, the Public Representative believes that imputation of the Collection Data Set is not justified and not needed to accurately estimate street time attributable costs.

Lastly, in order to provide a thorough analysis, a PIO analysis is conducted.[53] Despite the imputation modifications and the corresponding refined sample, the resulting 4 ZIP Code-days with Cook's D statistics greater than 0.1 in this scenario are the same as those mentioned in the Bradley Study, although with slightly larger values. *See* Cook's D – Bradley; Cook's D – PI. Consequently, following the same thought process as that in the Bradley Study, there is not sufficient reason to exclude these data points from the dataset.

Table 3: Comparison of CRE Model – Bradley Study and CRE Model – Partial Imputation

Marginal Times

Volume	CRE Model – Bradley Study	CRE Model – Partial Imputation	Percent Change
DPS	1.43	1.45	1.4%
Cased	0.86	0.86	0.0%
Sequenced	1.93	1.95	1.0%
FSS	2.98	3.08	3.4%
In-Receptacle Parcels	19.51	18.67	-4.3%
Deviation Parcels/ Accts.	64.26	65.92	2.6%
Collection	7.12	7.41	4.1%

[53] Please refer to PI Public Analysis, SAS output file "*Investigate PIOs with High.CooksD_NCVI - Modified.htm*" (Cook's D – PI).

Variabilities

Volume	CRE Model – Bradley Study	CRE Model – Partial Imputation
DPS	7.7%	7.8%
Cased	1.4%	1.4%
Sequenced	1.2%	1.2%
FSS	0.6%	0.6%
In-Receptacle Parcels	3.0%	2.9%
Deviation Parcels/ Accts.	7.3%	7.5%
Collection	1.2%	1.3%

Sources: CRE Model Combined Restricted Quad With Time Effects.lst and CRE Model Combined Restricted Quad With Time Effects_NCVI – Modified.htm

ii. Imputation is Unnecessary in the Delivery Data Set

The Postal Service also justifies imputation in the Delivery Data Set to preserve data. The Public Representative, however, does not believe this to be necessary as the Delivery Data Set is already large enough that imputation is not needed to maintain its large size and provide statistically significant results. Consequently, the second portion of the Public Representative's comparison analysis consists of modifying the imputation process discussed in section IV.a.3. to exclude any imputation. This involves eliminating imputation on the Collection Data Set nor on the Delivery Data Set. Rather, any observations that experienced imputation in either of these datasets in the Bradley Study are removed from their respective datasets instead of being imputed. [54] The Public Representative will refer to this modified methodology as the No Imputation (NI)

[54] In lieu of starting from scratch, the programs altered in section V.c.2.i. are also used in this modification along with additional modified programs which are included in the attached Library Reference PR-LR-RM2022-3-NP1, (NI Non-Public Analysis) folder "No Imputation," Directory 1 (NI Non-Public Directory 1) and Directory 2 (NI Non-Public Directory 2). Library Reference NI Non-Public Analysis contains non-public programs which remove the imputation procedure from the Collection Data Set and the Delivery Data Set preparation. The attached Library Reference PR-LR-RM2022-3-1, (NI Public Analysis) folder "No Imputation," Directory 2 (NI Public Directory 2) contains public programs which estimate variabilities in otherwise the same fashion as Proposal One.

analysis.[55] In contrast to the Bradley Final Data Set which contains 70,056 ZIP Code-days, the NI Final Data Set contains 62,640 ZIP Code-days. *See* Bradley Final Data Set; NI Final Data Set. The CRE model with time effects will again be the focus of this analysis.

The NI Final Data Set's summary statistics are about the same as those of the Bradley Final Data Set. Mean daily street hours per ZIP Code is about 93.82 hours and average percent of ZIP Code-days with FSS machines is about 19.7 percent in the Bradley Final Data Set, while mean daily street hours per ZIP Code is about 98.01 hours and average percent of ZIP Code-days with FSS machines is about 20.5 percent in the NI Final Data Set. *See* Table 2. Given their respective standard deviations, these averages are very similar. Additionally, Table 2 provides summary statistics for the corresponding CRE models relating to the Bradley Final Data Set and the NI Final Data Set. Specifically, it illustrates R^2 is 0.5137 and RMSE is about 671781.08 when the CRE model uses the Bradley Final Data Set, while R^2 is 0.5245 and RMSE is about 675423.06 when the CRE model uses the NI Final Data Set. The increase in R^2 by about 2.1 percent illustrates that NI, relative to Proposal One, materially improves the model's ability to explain the variation in street hours. The increase in RMSE by about 0.5 percent indicates the model's predictive ability reduces slightly. Note, the sample size decreased by about 10.6 percent when the modifications were applied, which could explain why the RMSE increased. This hinders any conclusions from being made based on the increase in RMSE, especially in the case of imputation where artificial data has been added.

[55] The resulting variabilities and marginal times from the NI analysis can be found at NI Public Analysis, SAS output file "*CRE Model Combined Restricted Quad With Time Effects_NI.htm*" (CRE Model – NI). The resulting finalized dataset can be found at NI Non-Public Directory 2, folder "Data," SAS data file "*fullzip_panel_NI.sas7bdat*" (NI Final Data Set).

After regressing the CRE model with time effects, it is clear the NI Analysis' results do not differ much from those of the Bradley Study.[56] Other than minor differences, most of the variables share the same sign and statistical significance / insignificance as their counterparts in the Bradley Study. *Id.* Also, as can be seen in Table 4, the CRE models for the Bradley Final Data Set and the NI Final Data Set result in very similar marginal times and variabilities. Although In-Receptacle Parcels' Marginal Time differs by 1.86 seconds between the two models, it is believed there is an overwhelming amount of support in favor of no imputation. Additionally, the variabilities are very similar amongst the two models.

In terms of model fit, the increase in R^2 when removing imputation from the data cleaning process seems to indicate that removing imputation improves model fit. In terms of impact, imputation overall does not have a large effect on the estimation of marginal times and variabilities. Thus, for theoretical and empirical reasons, the Public Representative concludes no imputation is the method preferred to Proposal One.

Lastly, in order to provide a thorough analysis, a PIO analysis is conducted.[57] Despite the imputation modifications and the corresponding refined sample, the resulting 4 ZIP Code-days with Cook's D statistics greater than 0.1 in this scenario are the same as those mentioned in the Bradley Study, although with values closer to one another.

[56] Please refer to CRE Model – Bradley and CRE Model – NI. Like with the PI modification, the intercept is no longer statistically significant at the 0.0001 level as its p-value is 0.0001. However, as justified in that section, the intercept is believed to be statistically significant which is in agreement with the result from the Bradley Study. With the exception of that just mentioned, dum37 (representing 07/13/2019) having a different sign, and dum13 (representing 03/23/2019) having a visibly different magnitude, the differences between the two sets of results are miniscule. Note since dum37 and dum13's coefficients are not statistically significant, these exceptions do not have any effective meaning.

[57] Please refer to NI Public Analysis, SAS output file "*Investigate PIOs with High.CooksD_NI - Modified.htm*" (Cook's D – NI).

See Cook's D – Bradley; Cook's D – NI. Consequently, following the same line of logic as in the Bradley Study, there is not sufficient reason to exclude these data points from the dataset.

Table 4: Comparison of CRE Model – Bradley Study and CRE Model – No Imputation

Marginal Times

Volume	CRE Model – Bradley Study	CRE Model – No Imputation	Percent Change
DPS	1.43	1.41	-1.4%
Cased	0.86	0.79	-8.1%
Sequenced	1.93	1.92	-0.5%
FSS	2.98	3.07	3.0%
In-Receptacle Parcels	19.51	17.65	-9.5%
Deviation Parcels/ Accts.	64.26	64.46	0.3%
Collection	7.12	6.99	-1.8%

Variabilities

Volume	CRE Model – Bradley Study	CRE Model – No Imputation
DPS	7.7%	7.7%
Cased	1.4%	1.3%
Sequenced	1.2%	1.2%
FSS	0.6%	0.6%
In-Receptacle Parcels	3.0%	2.7%
Deviation Parcels/ Accts.	7.3%	7.3%
Collection	1.2%	1.2%

Sources: CRE Model Combined Restricted Quad With Time Effects.lst and CRE Model Combined Restricted Quad With Time Effects_NI – Modified.htm

iii. Partial Imputation versus No Imputation

As mentioned previously, the PI modification results in a dataset with 64,440 ZIP Code-days while the NI modification results in a dataset with 62,640 ZIP Code-days. *See* PI Final Data Set; NI Final Data Set. As seen in Table 2, the two datasets have very similar descriptive statistics. Importantly, Table 2 indicates the resulting CRE model from the NI Final Data Set results in a larger R^2 (by about 0.2 percent) and a smaller RMSE (by about 0.4 percent) when compared to the CRE model run on the PI Final Data Set. Consequently, it is clear no imputation improves the model's ability to estimate street hours, although modestly. It is essential to note RMSE may be swayed by the varying sample sizes. Nevertheless, it is clear no imputation provides, although ever so slightly, an improvement in the model's predictive ability.

Comparing the respective coefficients between the CRE models shows that the overall results do not differ by much.[58] Other than minor differences, most of the coefficients agree in statistical significance/insignificance and sign. *Id*. Table 5 below illustrates marginal times and variabilities for the two models are quite similar. Marginally, however, no imputation appears to provide a slightly improved model as both marginal times and variabilities are slightly improved. Although these relative differences may be minor, this along with the increase in R^2 and the decrease in RMSE lead to the conclusion that no imputation is preferred to partial imputation and, thus, the overall preferred data cleaning process is that of no imputation.

[58] Please refer to CRE Model – PI and CRE Model – NI. Noticeable contrasts include dum37 (representing 07/13/2019) having differing signs in addition to dum13 (representing 03/23/2019) and dum25 (representing 05/18/2019) being visibly larger in the Partial Imputation model. Do note, however, that these coefficients are not statistically significant and, thus, their differences do not have material importance.

Lastly, in order to provide a comprehensive analysis, a PIO analysis is conducted. Similar to prior PIO sections, both datasets result in the same 4 observations being PIOs as those mentioned in the Bradley Study, although with Cook's D statistics closer to one another for the No Imputation model. *See* Cook's D – PI; Cook's D – NI. Since this aligns with that which occurred in section V.c.2.ii., there is insufficient reason to exclude these data points from the dataset.

Table 5: Comparison of CRE Model – Partial Imputation and CRE Model – No Imputation

Marginal Times

Volume	CRE Model – Partial Imputation	CRE Model – No Imputation	Percent Change
DPS	1.45	1.41	-2.8%
Cased	0.86	0.79	-8.1%
Sequenced	1.95	1.92	-1.5%
FSS	3.08	3.07	-0.3%
In-Receptacle Parcels	18.67	17.65	-5.5%
Deviation Parcels/ Accts.	65.92	64.46	-2.2%
Collection	7.41	6.99	-5.7%

Variabilities

Volume	CRE Model – Partial Imputation	CRE Model – No Imputation
DPS	7.8%	7.7%
Cased	1.4%	1.3%
Sequenced	1.2%	1.2%
FSS	0.6%	0.6%
In-Receptacle Parcels	2.9%	2.7%

Deviation Parcels/ Accts.	7.5%	7.3%
Collection	1.3%	1.2%

Sources: CRE Model Combined Restricted Quad With Time Effects_NCVI – Modified.htm and CRE Model Combined Restricted Quad With Time Effects_NI – Modified.htm

3. Practical Concerns

The imputation procedure is lengthy and will require additional time for the Postal Service to replicate in future updates to street time variability. Foregoing the imputation procedure would save time and money. The Postal Service notes that the current imputation procedure did not involve "excessive complexity or cost." Response to CHIR No. 5, question 6. It would be helpful if the Postal Service could provide the potential cost savings from not going through the imputation procedure, so that the Commission can gauge whether such a procedure is "excessively" costly. In addition, the Public Representative disagrees in the sense that the imputation procedures described in Proposal One are relatively complex in that they are applied with arbitrary thresholds to various different sets of observations. *See* Section IV.c.3. This complexity creates incentives for the Postal Service to modify its arbitrary assumptions and cutoffs for elimination and imputation in order to data mine, run many iterations of its models, and achieve more optimal street time variabilities in future rulemakings. *See* Response to CHIR No. 4, question 18. For example, the Postal Service's basis for imputing in certain situations was "where there were sufficient data to support imputation and the number of zero volume days was reasonable." Response to CHIR No. 5, question 6.b. As for the first condition, the Public Representative argues that there was not sufficient data to support imputation in the Delivery Data Set and, especially, the Collection Data Set in Section V.c.2. The second condition is when the Postal Service argues that imputation is "reasonable." This condition is too arbitrary and it gives the Postal Service an incentive

to rig imputation in future variability studies in order to achieve favorable variabilities without materially increasing the internal validity and / or robustness of the Top-Down Model.

It appears that the main rationale for imputing data is to preserve data for the sake of mitigating multicollinearity concerns. The Public Representative agrees with the Commission that multicollinearity is primarily a data problem. The analysis of the NI and PI models supports the conclusion that the expanded scope of the data allows the Postal Service to estimate reliable, statistically significant volume coefficients underlying variabilities without imputation in either the Collection Data Set or the Delivery Data Set.

The Public Representative believes that Proposal One should be approved by the Commission without imputation in the Collection Data Set or Delivery Data Set (as is conducted in the NI analysis). First, the Public Representative believes imputation is not justified conceptually given the limited amount of sample data the Postal Service has to make imputation estimates, especially for the Collection Data Set. Second, the Public Representative's analysis shows that estimating the Top-Down Model without any imputation improves statistical fit while not exhibiting deeper issues of multicollinearity, namely the lost significance of primary explanatory variables. Importantly, foregoing imputation results in practically similar estimates of variability and marginal times. Third, the Postal Service can save both time and money, increase transparency, and avoid incentives to data mine its future variability results by foregoing the imputation strategy detailed in Proposal One.

d. Special Studies of Collection Volumes or Development of Daily Collection Volume Should Be Expected for Future Variability Updates

The Postal Service suggests that "the small size of the bias [from omitting collection volumes in its analysis] suggests it may be possible to update the street time variability equation in the future without having to field a special collection volume study." Response to CHIR No. 2, question 4. While the Public Representative agrees that the size of the bias when omitting collection volumes in a model of street time variability appears small, the Public Representative urges the Commission not to rely on the limited Collection Data Set for an extended period of time. The fact is that little seems to be known about trends and seasonality in collections from customers' receptacles, and the last study of collection volumes reviewed by the Commission occurred in FY 2013. The Postal Service argument that this collection data from a two-week period from January and February is a good proxy for collection data for the rest of the year is not based on the merits of the presented data. It states that "the average day-of-week collection volume from the two-week study turns out to be the best available proxy for collection volume (obtained at customers' receptacles) throughout the year, as it is the only such data available." Response to CHIR No. 5, question 4.a. Unlike in Docket No. RM2015-7, where the Postal Service rebutted a seasonality concern raised by the Public Representative (of that docket) by stating that the parcel and accountable special study occurred neither "at the peak nor the trough of parcel volumes," in the instant case, the Postal Service appears to have no information relating to collection mail seasonality. *See* Response to CHIR No. 5, question 4.a.

The Public Representative recognizes that in Order No. 2792, the Commission accepted the results from special studies underlying Proposal Thirteen, including a similar special collection study, but in that case, it continued to urge the Postal Service to develop

reliable collection volumes. It should do the same in the instant case. The Commission noted that "reliable daily volumes of collection mail" may eliminate the need for "extensive and time-consuming special data collection studies." Order No. 2792 at 65. Daily collection volumes have been presented in this docket, but there is still much uncertainty relating to the nature of this data and the Public Representative would argue this data cannot be used reliably (as a sole source) for future rulemakings. The collection data presented may be appropriate for the purposes of the instant docket in updating street time variabilities, but it should not be the only source of data relating to collection volumes relied upon for future analyses.

e. Modified Proposal One Impact

The impact of Proposal One is a decrease of street time attributable costs by $2.178 billion. *See* Section IV.c.3. The estimated variabilities do not change materially in the Public Representative's NI analysis. See Section V.c.2.ii. While the Public Representative has not explicitly calculated the difference in the impact if no imputation were performed, the similar top-down model results of an analysis with no imputation suggests that the impact should be similar to those identified in the Bradley Study. Generally, the variabilities and costs for market dominant products should decrease and the variabilities and costs for parcels should increase compared to the present methodology. Before its decision, the Public Representative urges the Commission to ask the Postal Service for the restated impact figures for a variability study without Collection Data Set imputations and Delivery Data Set imputations or, if the Commission prefers, a study just without Collection Data Set imputations.

f. Unobserved Heterogeneity

The Postal Service submits a new estimation procedure called a correlated random effects model to perform the variability analysis and estimate variabilities for street time products. The Public Representative agrees with the Postal Service that unobserved heterogeneity is a serious concern in variability analyses, in general, and with street time variability, in particular. For example, in Docket No. RM2020-13, the Commission identified and rejected the Postal Service's proposal for estimating mail processing variabilities because a fixed-effects model alone was insufficient in controlling for unobserved heterogeneity in mail processing time between facilities. In the case of street time, the Postal Service demonstrates significant variation in street hours for ZIP Codes with similar volumes. *See* Bradley Study, Figure 11. The correlated random effects model confirms the presence of unobserved heterogeneity. Therefore, it is important that Proposal One address this unobserved heterogeneity and the Commission's previously stated concerns with fixed-effects models. The Postal Service's CRE model appears to satisfy these conditions, but the Public Representative admits the underlying econometrics exceed his current realm of expertise. *See* Bradley Study at 86-89. Various CHIRs have been issued in this proceeding relating to the appropriateness of the CRE model in this context, and the Public Representative reserves the right to comment on this issue further, if necessary, in reply comments. *See* Response to CHIR No. 1, question 3; Response to CHIR No. 2, questions 1, 2; Response to CHIR No. 4, questions 5, 6; Response to CHIR No. 5, questions 3, 8, 9, 10. Relatedly, the Postal Service's decision to restrict the number of estimated interaction terms in the CRE model also makes sense in order to alleviate potential multicollinearity concerns and conforms to Commission precedent. *See* Section IV.c.1.

VI. Conclusion

The Public Representative recognizes the careful and thoughtful analysis submitted by the Postal Service in Proposal One. The methodology of Proposal One in estimating of estimated variabilities and marginal times are an improvement over the current methodology in 39 C.F.R. § 3050.11(a). The results are more accurate because they rely on a greatly expanded dataset of delivery volume, collection volume, and Non-Postal data, and the Top-Down Model has been refined to explicitly control for unobserved heterogeneity in street time. The results are more complete because there is a more complete set of underlying data, and the Top-Down Model can simultaneously estimate variabilities for all of the street time products. The quality of the data has also been improved because more data are taken and merged from operational Postal Service data systems, which have been refined and approved by the Commission over the past several years. While recognizing these improvements, Proposal One could be improved further if no imputation procedures were conducted because they are conceptually and statistically unfounded, do not significantly change the model estimates, and are costly. Finally, the Postal Service should continue to develop its understanding of seasonal and year-over-year trends in collection volumes.

The Public Representative respectfully submits the foregoing comments for the Commission's consideration.

Respectfully submitted,

Philip Thomas Abraham
Public Representative

901 New York Avenue, NW, Suite 200
Washington, DC 20268-0001
Office Phone: 202-789-6825
Email: Philip.abraham@prc.gov