White Paper:

Lessons from Rebuilding the
FCC's Quantile Regression Analysis

By Vincent H. Wiemer and Michael J. Balhoff, CFA

Alexicon Consulting *Balhoff & Williams, LLC*

February 2013

Table of Contents

White Paper

Lessons from Rebuilding the FCC's Quantile Regression Analysis

By Vincent H. Wiemer and Michael J. Balhoff, CFA
Alexicon Consulting and Balhoff & Williams, LLC
February 2013

Executive Summary

This White Paper evaluates the Federal Communication Commission's ("FCC") quantile regression analysis ("QRA") model, revised on April 25, 2012 and initially implemented on July 1, 2012.[1] The QRA benchmarks the costs recoverable from High Cost Loop ("HCL") Universal Service funding ("USF") for small rural carriers across the United States on the basis of cost comparisons among the majority of small rate-of-return local incumbent telephone companies.[2]

Our White Paper is prompted by the significant criticism over the last year concerning the QRA's overall design and execution, the concern over the lack of transparency related to data underlying the QRA's input variables, and the potential that the QRA model could fail the statutory purpose found in Section 254 of the Telecommunications Act of 1996.

- **Purpose.** This White Paper seeks to provide a disciplined review of the QRA, which is being used to automatically and presumptively disallow cost recovery above a benchmark determined by the highest 10% of costs of the companies in the rural rate-of-return industry. Prior to the QRA, those costs were generally recoverable by carriers meeting federally-defined universal service objectives.[3] The QRA model then reapportions the "disallowed" funding to other rural carriers.
- **Approach.** This White Paper documents the authors' reconstruction and analysis of the QRA.[4] The authors reproduced the development of the QRA's data inputs and explored the effect of

---

[1] The FCC released results for a new application of the QRA on January 30, 2013, using the same coefficients developed in April 2012; the results now reduce support for about 160 carriers by contrast with the April 2012 version that reduced recovery for about 100 carriers; the spreadsheet with the 2013 results is available at http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-318654A1.xlsx.

[2] According to National Exchange Carrier Association 2011 High Cost Loop data there were 1,139 rural rate-of-return study areas at the time the QRA was implemented. Of these, 726 study areas were used in the cost comparison, 10 were not used because of data issues, 28 were acquired exchanges, and 369 were average schedule companies for which costs were not available.

[3] The QRA provides for a phase-in of the initial disallowance over eighteen months, and then the full effect of the 10% disallowance will be effective from 2014; subsequent years will not be phased in, as the full 10% disallowance will be effective each of those years.

[4] The authors are aware of no other disciplined review of the QRA to date. Vincent Wiemer is a principal at Alexicon Consulting (founded in 2002), which provides full management resources, including cost consultant services to the rural telecommunications industry; Michael J. Balhoff headed Legg Mason's technology and telecommunications equity research group for 16 years before founding a financial and regulatory consulting firm,

those inputs on the overall performance of the model. This report describes the insights derived from rebuilding the model. Those findings are compiled under four headings.

- *Review of the QRA.* The model is reviewed initially by analyzing the individual variables and the data sources that the FCC used in creating coefficients for the QRA.
- *Perspective on the QRA's role in meeting policy obligations.* The QRA is a tool that is designed to accomplish a congressionally-mandated and affirmative policy goal, which is universal service. The second section responds to the question about whether the tool, as currently formulated or as potentially revised, serves the larger purpose. The performance of the model is evaluated against the statutory obligations which are spelled out as "specific, predictable and sufficient Federal and State mechanisms to preserve and advance universal service."[5]
- *Recommendations to improve the QRA.* We offer suggestions about how the model might be improved, if a model is to be used at all. Recommendations flow from the findings described in the first section.
- *Appendices.* Major elements of the detailed analyses are found in the appendices of this White Paper. Additionally, there is an appendix that analyzes certain aspects of the FCC's commentaries about the model, relying on FCC assertions in the two Orders.

- **Summary insights.** The QRA is a complex tool and merits a careful review because of the importance of Universal Service to customers, economies, communities, and national broadband policy. The simple summary of our findings is that the QRA could be a valuable mechanism—with revisions to apparent flaws—if the tool is used in identifying *potentially* excessive costs, but the QRA is currently and likely always will be inaccurate to such an extent that some human judgment should be used in evaluating whether a carrier's costs are in fact excessive.

The specific insights into the current QRA—all related to the most recent QRA as ordered in April 2012—can be summarized as follows.

- There are specific problems with the design and execution of the QRA.
  - *Valid cost relationships cannot be determined by the QRA because the costs are improperly defined.* Since the fundamental goal of regression analysis is to establish relationships between costs (dependent variables or outputs) and cost-predictors (independent variables), the definition of the inputs and outputs to be computed is critical to the success of the statistical model. However, the QRA uses ambiguous definitions that vary from one case to the next. So, for example, plant depreciation expense (used significantly to define capital expenditures) is arbitrary and inconsistent, depending on whether a state regulatory body sets the depreciation rate at, for example, 4% or 7% or 10%. So the computation process is comparing target costs that are not comparable across the universe of companies. The result is that the model, based on non-comparable variables, fails to predict costs reliably. In addition, the QRA focuses on an allocation

Balhoff & Williams, LLC in 2004 and an investment banking firm, Charlesmead Advisors, LLC in 2011; more detailed biographical information for the authors is included in the appendices.

[5] Telecommunications Act of 1996, Section 254(b)(5).

scheme that artificially separates operating ("opex") and capital expenditure ("capex") expenses, in spite of the fact that certain costs could be incurred in capex that could reduce opex or vice versa.[6] The FCC is rumored to be evaluating this problem, the resolution of which would improve the model but not resolve the wider concerns raised here.

- *Of the sixteen independent variables used in the QRA, fourteen appear to have meaningful problems.* The notable problems include (i) the use of inaccurate or outdated data in the source databases; (ii) questionable or clearly flawed assumptions; (iii) weak or no cost causation which make the use of certain variables problematic in a predictive model; (iv) obvious errors in the results the variables generate; (v) too few source-data points for statistical reliance; and (vi) obviously low predictive values.[7]
- *Poor procedures related to statistical data analysis significantly weaken the results.* The regression analysis (i) provides poor overall predictive capability; (ii) fails to examine the effects of outliers that may distort the calculations;[8] (iii) does not assure comparability of costs between individual companies; (iv) does not identify and correct obviously incorrect results (as noted above); and (v) omits important cost-causative realities from the analysis.
- *Conceptual issues make the achievement of stated policy goals unlikely.* The design appears to frustrate the political goal of incenting broadband deployment as the statistical tool compares costs at companies that have little broadband infrastructure with those that have more broadband infrastructure, and is more likely to disallow the recent and relatively more costly broadband investments, particularly in a transition from old plant to new plant at a time when costs are reasonably higher.

- *If a model is used, a better-functioning model is possible.* The third section of this White Paper proposes constructive recommendations in light of the findings in the first section. These suggestions include investigations about (i) regressing for a single set of costs rather than the two outputs—opex and capex—used in the current regression; (ii) establishment of valid cost relationships; (iii) the proper examination of outliers; (iv) the exclusion of costs, such as taxes, over which the companies have no control from the benchmarking of "excessive" costs; and (v) evaluation and remedy of certain unintended

---

[6] We understand that the FCC is currently considering how to combine capex and opex costs to mitigate this specific problem in allocating costs.

[7] USF/ICC Transformation Order, Appendix H, para. 22 states that the model included variables that were not significant in all the regressions; "Not all the variables are significant in each regression, and there are some variables (such as the log of land area in urbanized clusters) that are not significant in any of the regressions. We chose to use all the variables in all the regressions so long as the parent variable (such as land area) had at least one child variable (such as land area in a non-urbanized area) that was significant for at least one of the regressions in the analysis. While this meant that some regressions had many insignificant variables, this was not a problem because the goal of the regression was not to determine statistically significant correlations, but instead to generate 90th percentile predictions, which are unaffected by the addition of insignificant variables."

[8] USF/ICC Transformation Order, Appendix H, para. 10 contends that outliers are not significant in the model because of the use of a quantile regression. We address this contention more fully in *Appendix E*.

consequences. Such a model could "improve" the predictive capacity compared with the current QRA model.

- *The use of a QRA to presumptively disallow funding falls short of good policy.* The statutory goal is to provide specific, predictable and sufficient funding to preserve and advance Universal Service. In the 2001 reform of USF, the FCC relied on data studies sponsored by the Federal-State Joint Board on Universal Service ("Joint Board"). Those studies demonstrated that there was a significant difference in costs, not only *between* rural and non-rural companies, but also *among* rural-centric companies. The task force reported that modeled costs in rural regions were insufficient in predicting appropriate investment and operating costs for small and vulnerable carriers. That finding of "substantial diversity" from one rural region to another led the Joint Board to recommend—and the FCC to mandate—rural funding in the 2001 USF reform on the basis of actual costs, using a modified embedded cost approach.[9] To the best of our knowledge, no studies or data have overturned that conclusion. The first two sections of this White Paper reinforce the insights from a decade ago concerning the problems with predicting a company's costs based on so diverse a set of companies. The FCC's new QRA uses a model that, by the FCC's admission, achieves about 62%-67% accuracy, based on the correlations between the model's independent variables and the outputs.[10] The FCC's purpose in using the QRA is to create accountability and discipline with respect to funding by automatically disallowing approximately $94 million once the plan is fully implemented. The automatic reallocation of the presumed "excessive funding" out of total high cost funding of $4.5 billion is approximately 2% of the total, which is a small portion of the fund but could have a significant effect on individual small carriers where the costs may actually be appropriate. The issue is that presumptive disallowance based on an approximate calculation process is likely to lead to harmful outcomes without ongoing consideration of whether the model has found or failed to find excesses.

---

[9] Rural Task Force, *The Rural Difference: White Paper 2,* January 2000, pp. 5-6; available at http://www.wutc.wa.gov/rtf/old/RTFPub_Backup20051020.nsf/?OpenDatabase. The Rural Task Force was created by the Joint Board on Universal Service to study potential reforms; its appointed membership included a wide range of industry interests and experts: Chairman William R. Gillis, Commissioner, Washington Utilities and Transportation Commission; Robert Schoonmaker, Vice President, GVNW Consulting, Inc.; Thomas Beard, President, National Phone Company; Carol Ann Bischoff, Executive Vice President and General Counsel, Competitive Telecommunications Association; Jack Brown, Management Consultant Golden West Telecommunications Cooperative, Inc.; David R. Conn, Vice President Law and Regulatory Affairs, McLeod USA, Inc.; Gene DeJordy, Executive Director: Regulatory Affairs, Western Wireless Corp.; Billy Jack Gregg, Director, West Virginia Consumer Advocate Division; Joel Lubin, Regulatory VP-Law and Public Policy, AT&T; Joan Mandeville, Assistant Manager, Blackfoot Telephone Company; Christopher McLean, Deputy Administrator, Rural Utilities Service, USDA; Gwen Moore, President, GEM Communications; Jack Rhyner, President and CEO, Telalaska; Jack Rose; David Sharp, President and CEO, Virgin Islands Telephone Corp.; Stephen G. Ward, Public Advocate, State of Maine Public Advocate Office. The RTF relied upon the professional support services of the National Exchange Carrier Association; The National Telecommunications and Information Administration--U.S. Department of Commerce; The Rural Utility Service--U.S. Department of Agriculture and The Rural Policy Research Institute and the University of Missouri Office of Social and Economic Data Analysis.

[10] The FCC has reported that the revised QRA has a pseudo-R-squared, which quantifies a confidence level, of 0.62 for the opex per line and 0.67 for the capex per line, where a value of 1.0 reflects the highest predictive value, at least based on the statistical calculations related to the input data.

- *Recommendations.* The FCC's use of a model to create accountability may be a sound concept. However, the use of an unproven and demonstrably inaccurate model to *automatically* exclude funding appears to run contrary to the statutory mandate for specific, predictable and sufficient funding. It appears that accountability for use of the universal service funds and overall fund size were ordered by the FCC Commissioners, but the Commission did not order the details of how the QRA is to be shaped or used. Accordingly, the authors recommend changes in light of the major problems with the current model and the still-insufficiently predictive results of an improved model, as well as the relatively small benefit generated through the automatic cuts.
  - The model can be improved. The model should be revised to incorporate the suggested improvements in the third section of this White Paper. Specifically, improved data sources and variables should be identified and used; the opex and capex calculations should be combined into a single benchmark; and costs that cannot be controlled by the carrier—because of exogenous factors that vary widely from one jurisdiction to the next including tax and depreciation rates set by governmental authorities—should not be used to "identify" excessive and unrecoverable expenses.
  - The model should trigger reviews rather than automatic disallowances. The authors are not recommending that the model be discarded, but that it should be used as a tool in guiding policymakers toward sound policy decisions rather than as the final determinant of policy decisions. The model should be used to identify *potentially* excessive costs, and it should precipitate a review of a carrier's costs to determine whether those costs are in fact excessive. The examinations could occur at the FCC or at the state commissions or at some other designated entity such as the Universal Service Administrative Company, thereby achieving greater accountability. Importantly, such a review will better assure that there is no undue harm to customers, local economies, employees, or providers of capital that support universal service goals. Additionally, the use of the model to precipitate reviews should create better ongoing information for regulators and for companies, as regulators have a mechanism that prompts analysis and monitoring of the QRA as well as assurance that there is specific and sufficient universal service funding. Importantly, a review process will give carriers greater confidence that they can invest and will not be subjected to unpredictable and insufficient investment cycles, and the process will make it more likely that the investment-based goal of universal service is realized.
  - An advisory committee should be established. Finally, it is recommended that an interdisciplinary committee should be established to advise the state commissions or the FCC or the designated review entity. The advisory committee could evaluate the preliminary findings, thereby triggering a review to determine if costs are actually excessive, so the committee of the FCC could then clarify what funds should be reallocated to other carriers. In this way, the QRA becomes a tool to further Congress' and the FCC's purpose in achieving ubiquitous broadband deployment that serves the goal of Universal Service.

Introduction

The FCC ordered sweeping reforms of universal service and intercarrier compensation ("ICC") in late 2011. The USF/ICC Transformation Order was released on November 18, 2011.[11] In that reform, the FCC introduced the use of a regression analysis to limit the amount of cost recovery available to certain individual rural rate-of-return carriers. Significant commentary and criticism of the tool precipitated a modification of the QRA in April 2012.[12]

On April 25, 2012, the FCC's Wireline Competition Bureau ("the Bureau") released an Order DA12-646, *In the Matter of Connect America Fund High Cost Universal Service Support* ("HCLS Benchmark Order"). The HCLS Benchmark Order responded to many of the criticisms related to the original proposed tool introduced in November 2011 in the USF/ICC Transformation Order.[13] The HCLS Benchmark Order detailed a revised QRA that would be used to create "benchmarks" which allowed for capping the amount of funding for capital and operating costs recovered through the High Cost Loop Fund ("HCL"), one of the fund-elements of the USF program; the recaptured funds were then to be reallocated to other carriers.[14] Minor updates will occur in 2013, but the coefficients for the QRA are to be recomputed annually beginning in January 2014.[15] The FCC did release the 2013 computations at the end of January 2013, with an apparent negative impact on 60% more of the carriers (about 160 carriers were to incur reductions by contrast with about 100 in 2012); we understand that certain errors may have occurred in the calculations and the FCC might issue a further revision.[16]

In May 2012, in an effort to understand the operation and potential problems with the QRA, several independent local exchange carriers sent letters to the Bureau requesting the workpapers and files that had been used to develop the dependent and independent variables in the QRA.[17] The carriers' contention was that it was impossible to respond to the model if the carriers' experts did not have access to the underlying data and to the FCC's assumptions that affected how those data were used in the QRA. Initially, the Bureau provided only a list of the databases and summary documentation of the processes

---

[11] *See Connect America Fund; A National Broadband Plan for Our Future; Establishing Just and Reasonable Rates for Local Exchange Carriers; High-Cost Universal Service Support; Developing a Unified Intercarrier Compensation Regime; Federal-State Joint Board on Universal Service; Lifeline and Link-Up; Universal Service Reform—Mobility Fund*; WC Docket Nos. 10-90, 07-135, 05-337, 03-109, CC Docket Nos. 01-92, 96-45, GN Docket No. 09-51, WT Docket No. 10-208, Report and Order and Further Notice of Proposed Rulemaking, 26 FCC 17663 (2011) (*USF/ICC Transformation Order and FNPRM*); (hereafter "USF/ICC Transformation Order).

[12] *See, e.g., The FCC's Quantile Regression Analysis is Fatally Flawed, Period : Commenters Provide Dozens of Arguments Against QR, None in Favor* (JSI Capital Advisors, February 7, 2012), available at http://www.jsicapitaladvisors.com/monitors/2012/2/7/the-fccs-quantile-regression-analysis-is-fatally-flawed-peri.html.

[13] For example, the USF/ICC Transformation Order at Appendix H of the FNPRM had proposed to create eleven caps (four capex caps and seven opex caps), which were modified in the HCLS Benchmark Order to include only two, one for opex and the other for capex.

[14] USF/ICC Transformation Order, para. 214.

[15] *Id.*

[16] FCC, "Appendix: Quantile Regression Cost Per Loop (CPL)," January 29, 2013, posted February 4, 2013, available at http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-318654A1.xlsx; see also http://apps.fcc.gov/ecfs/comment/view?id=6017160978.

[17] *See, e.g.,* Letter to Sharon Gillett, FCC, May 1, 2012, available at http://apps.fcc.gov/ecfs/document/view;jsessionid=72snPjBbppV2mmgwJxLbFVWGcrHpLT6Ly99hWygLkL2gFm1x22T4!-1969853125!-1221852939?id=7021915584.

used to develop the QRA variables. Finally, on June 18, 2012, less than two weeks before the implementation of the QRA, the Bureau posted more detailed information about certain elements that explained assumptions and code related to the QRA.[18]

Since June 2012, a team of people under the direction of Vincent Wiemer of Alexicon Consulting, in consultation with Balhoff & Williams, has worked to reconstruct the FCC's model. The plan was to recreate the model, then test whether it generated results similar to those reported by the FCC, and, finally, disaggregate the variables and the data to understand the critical drivers of the QRA. It was hoped that such a process would permit all stakeholders to gain visibility into the model in a way better than what was previously available. It was also hoped that other industry experts might be able to use the study to affirm the value of certain elements, propose changes that would improve the model, and provide sound insights to policymakers at the FCC, state commissions, and Congress.

SUMMARY OF THE QRA

The purpose of the FCC's statistical methodology is "to identify study areas that have capex [capital expenditure expense] and opex [operating expense] costs that are much higher than their similarly-situated peers and to cap their cost recovery at amounts that are no higher than the vast majority of similarly-situated study areas."[19] To accomplish its goal, the FCC employed a quantile regression analysis to estimate the relationships between assumed company cost drivers and dependent outputs that, as of April 2012, were operating and capital costs.[20] The idea was to select certain factors that could affect changes in a carrier's costs and use those factors to "predict" appropriate costs.[21] The FCC concluded that capex and opex are affected by scale of operations, age of plant, customer dispersion, and geography. The sixteen independent variables used in the QRA were based on:

- ***Scale (4 independent variables)***: number of loops, road miles, number of road crossings, number of study areas under common control in the state;
- ***Age of Plant (1 independent variable)***: percentage of undepreciated plant;
- ***Customer Dispersion (3 independent variables)***: customer density, number of exchanges in the study area, percentage of households in urban areas;
- ***Geographic factors***: The FCC included construction difficulty and geographic regions under a single heading. It appears more logical to differentiate this heading.
    - ***Construction difficulty (3 independent variables)***: soil difficulty index; percentage of bedrock; and frost index ("climate")

---

[18] The FCC background information was posted on June 18, 2012 by the Bureau in a file named *Geospatial Data Creation Algorithms.zip,* available at http://transition.fcc.gov/wcb/iatd/neca.html.

[19] HCLS Benchmark Order, para 59.

[20] Study areas are service regions within a state. The FCC and the states require reports from those study areas to detail service, investment and other regulatory data. It is possible that an incumbent local telephone will have only one study area within a state. Generally because of acquisitions of other telephone companies, a carrier may have several study areas within a state, requiring two or even three reports that will go to the FCC or to the state commissions.

[21] HCLS Benchmark Order, para 81.

- ***Geography (5 independent variables)***: percentage of study area on tribal land, percentage of study area on national park land, and regional location (Alaska, Midwest, and Northeast).[22]

The definitions of capex and opex per study area were developed from the National Exchange Carrier Association ("NECA") High Cost Loop ("HCL") support algorithm.[23] The NECA HCL algorithm takes study area cost data and, through a series of allocations, distributes and develops the capital and operating costs used for the local loop network. Historically, the resulting study area cost per loop was compared to a National Average Cost per Loop ("NACPL") and then study areas with loop costs significantly in excess of the national average were eligible to receive support funds, based on defined formulae.

The QRA takes the capex and opex dependent variables and the sixteen independent variables applied to each of 726 rural study areas and attempts to estimate relationships among them. In simple terms, the QRA seeks to analyze whether changes in the independent variables related to number of loops, road miles, customer density, and other factors predict changes in the dependent variables which are capex and opex costs.

For example, the QRA has 726 data points of actual capex costs and 726 data points that represent the number of loops in those study areas. A linear regression is run to "determine" to what extent a change in the number of loops correlates to a change in the capex cost. A simple regression is based on an analysis with a single independent or predictor variable. Multiple regression analyses assess the relationships among multiple predictor (independent) variables and a dependent variable.

Regression outputs take the form of a mathematical equation that describes a straight line through the data. In most forms of linear regression, that line is the average of the data, with the linear equation representing the predicted results. Going back to the simple example of loops and capex cost, the linear regression would generate an equation that would allow inputting the number of loops to compute the average capex cost based on the 726 study areas (assuming that the number of loops is in fact a predictor of properly defined capex costs).

The use of a quantile regression, in contrast with a least squares regression, allows the lines to be drawn through the output data points at places other than the average. The lines can be drawn at "quantiles," a location where a percentage of the data observations fall above and below the line. For example, the median is the 50th quantile—the line where 50% of the data observations are above the line and 50% below the line. The FCC's selection of a 90th quantile benchmark means that the QRA outputs a formula that describes a line where 10% of the actual study area capex and opex costs are above the line and 90% below the line (see Figure 1). That same cost relationship formula can then be used to determine a point on the capex or opex line (the "capped cost") by inputting the values drawn from any individual

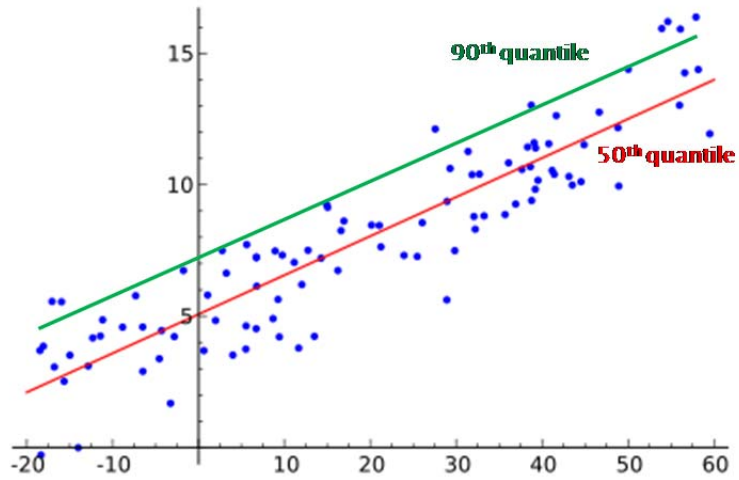
---

[22] The Bureau discarded geographic variables for the South and West because the variables did not appear significant.

[23] The FCC defined capex costs as the sum of depreciation expense and return on capital attributed to local loop cable and circuit equipment (NECA HCL algorithm steps 17, 18, 23 and 24). The FCC defined opex costs as the sum of plant maintenance, network and general support, network operations, corporate operations, operating taxes, benefits, and rents attributed to local loop cable and circuit equipment (NECA HCL algorithm steps 13, 14, 15, 16, 19, 20, 21 and 22).

study area set of predictor variables. The lesser of a study area's actual cost or capped cost is used in the calculation of high cost loop support.

Figure 1: Linear Regression Quantile Illustration



Source: Alexicon Consulting, 2013.

In summary, the FCC has chosen to use a quantile regression analysis as a tool to assess capital expenditure costs and operating costs that, according to the tool, identify the top 10% of the costs. The assumptions are that the tool properly defines the dependent costs, has appropriate predictor variables, employs accurate databases, and generates sufficiently reliable results. To simplify the regulatory oversight, the FCC has chosen to employ the tool in a way that presumptively disallows the recovery of certain investment and operating costs that fall above the 90th quantile. This approach is based, at the present, on no human review of whether the costs are *in fact* excessive in meeting the carrier's obligations to serve its customers.

Our study attempts to examine critical assumptions about the QRA variables, databases, and predictive capabilities of the tool. It is again our purpose to provide a more detailed fact study so that the review of the QRA can be better informed.

Section 1: Review of the Quantile Regression Analysis

This first section of the White Paper is a report on the authors' findings arising from rebuilding the FCC's QRA. Every effort is made to analyze without bias the statistical tool's effectiveness and to provide specific and constructive insights.

This initial section is organized into several parts which we summarize here.

- ***Methodology employed in reconstructing the QRA.*** The methodology of our analysis is summarized in this section and detailed in *Appendix A*. Our approach involved assessment of each of the individual variables as well as the underlying data sources and the Bureau's development processes.

- ***Specific Findings.*** The QRA relies in its final form on sixteen "independent" variables to generate predictions for two "dependent" variables, which are capex and opex costs. Our specific findings can be divided into (i) insights derived from studying the capex and opex costs (the dependent variables), (ii) observations about the independent variables, and (iii) other general findings.

    - *Findings - Costs.* The opex and capex costs, as defined in the model, contain NECA-modeled, allocated "expenses" that are not comparable across the universe of carriers, and the effect is to render the QRA's results unreliable when applied to real-world costs. The flawed definition of capital and operating costs seriously compromises the ability of the QRA to establish meaningful relationships between costs and cost predictor variables. Additionally, the artificial segregation of capex and opex ignores the economic reality of trade-offs between investment and ongoing expenses and likely penalizes some carriers that have made prudent decisions to reduce total costs; that is, opex may be relatively low because of higher capex investment or vice versa.[24]

    - *Findings - Independent Variables.* The QRA input variables have multiple and significant problems. Of the sixteen independent variables, fourteen of the QRA inputs contain one or more noteworthy errors. The notable problems include (i) the use of inaccurate or outdated data; (ii) questionable or demonstrably flawed assumptions; (iii) weak or no cost causation which make the use of such variables problematic in a predictive model; (iv) obvious errors in the results the variables generate; (v) too few source-data points for statistical reliance; and (vi) obviously low predictive values. The disappointing conclusion is that significant errors in data or in assumptions weaken the predictive capabilities or, more to the point, raise grave concerns about the use of the QRA as a tool to presumptively disallow individual carrier cost-recovery.

    - *Other General Findings.* The QRA omits important cost-causative factors that affect construction costs or operating expenditures, such as extreme weather conditions or conditions that require burying or protecting outside plant. The model also suffers from

[24] It is our understanding that the FCC may change the model to combine the capex and opex dependent variables to resolve this potential problem.

the use of data for independent variables that are not comparable among companies, such as widely divergent tax regimes over which the companies have no control. And the QRA fails to consider the problematic impact of outliers, in spite of the FCC's assertion that a quantile regression improves upon the problem of outliers over what occurs in a least squares regression.

Methodology employed in reconstructing the QRA

It is apparent that the FCC worked diligently to craft its reforms of USF and intercarrier compensation in 2011. It should also be affirmed that the task was complex and challenging, requiring many compromises.

With respect to the newly-introduced QRA, the challenge is particularly daunting because the FCC seeks to design a tool that analyzes construction and operating costs across regions and operations that are extraordinarily disparate. The variances include differing construction challenges because of terrain and weather, widely divergent stages of plant modernization, carriers with denser service areas or notably isolated customers, and remarkably different carrier sizes. *The question that arises is whether a regression model can predict capex and opex with sufficient accuracy to reconcile those differences, and whether such a model can be relied upon as the final determinant in identifying excessive expenditures.*

This initial sub-section describes in summary fashion how the QRA was rebuilt as we attempted to answer the questions outlined above. Again, *Appendix A* supplies the fuller explanation of the reconstruction work.

1. Audit the Independent Variables

Alexicon coordinated a deliberate process to review the QRA and the variables used to benchmark a carrier's operating expenses and capital expenses.

The reconstructive process began by focusing on the FCC's development of the independent variables used to predict capex and opex. Due to the specialized nature of the data and the programs involved, the team enlisted the expertise of EI Technologies, LLC, which is a geospatial information systems ("GIS") consulting firm.[25]

Compile the FCC's data resources. The team first identified the data resources, which were reported in Appendix A of the HCLS Benchmark Order and in the FCC process document that had been released on June 18, 2012. The databases included NECA High Cost Loop input data, ESRI ArcGIS StreetMap, Tom Tom Telecom Suite 2011.09, 2010 Census data, STATSGO2 database, the USDA Hardiness Index, Tribal Lands data, National Parks data, the number of local telecommunications Study Areas, and state and regional data from the FCC process documentation. Our team compiled the databases available from

---

[25] EI Technologies is a geospatial information systems ("GIS") consulting firm with over 20 years of experience in GIS Application Development and Programming and other GIS services. EI Technologies' lead applications developer has been recognized by the ESRI user community as an ESRI MVP (Most Valued Programmer). Its developers have in-depth knowledge of such technical and scientific topics as hydrology, geology, and civil engineering. The project was headed by Mr. Nirav Shah, President of EI Technologies, LLC. Mr. Shah is a Certified GIS Professional (GISP) and was named Industry Visionary by GeoWorld for seven straight years.

public sources, and entered into licensing agreements for the proprietary databases and software that were reported by the FCC to have been used in constructing the QRA.

Reproduce the development of independent variables. Under Alexicon's direction, EI Technologies, LLC, was assigned responsibility for reproducing the QRA independent variables based on the data sources described above. The use of data in the calculations involved evaluation of the FCC's assumptions and intermediate steps that were not well understood until the reconstruction occurred. We note that it would have been virtually impossible to reformulate the independent variable calculations without the FCC's release of process information in June 2012.

Validate the source data and processes used. EI Technologies focused on the initial assessment of the quality of independent-variable source data used, as well as identification of factors, problems, or other limitations arising from the processes used by the FCC.

Test the outcomes. The output data were tested to identify whether the team's processes approximately replicated the FCC's results. In the end, the reconstructed processes generated results virtually identical to those of the FCC.

2. Analysis of the QRA and the relationship of the individual variables

The next stage of the review focused on the analysis of how the QRA performs, with a goal to understand the interaction of the model's component parts in a way that was far more detailed than was possible before the FCC disclosed its processes on June 18, 2012. The initial analysis was concentrated on the independent variables (cost predictors) and dependent variables (capex and opex costs).

Reproduce the QRA's results. Alexicon compiled the information from EI Technologies and again reproduced the QRA formulated by the FCC using derived variable and cost information for all 726 carriers in the FCC's tables. The discipline included a recalculation of the capex and opex limits and resulting HCL support amounts for each carrier, so that the model could be verified once again against the FCC's reports for individual carriers.

Analyze the dependent variables. The general purpose of multiple regression is to learn more about the relationships between several independent (or predictor) variables and a dependent variable. The team analyzed the ability of the QRA to establish valid and meaningful cost relationships.

Analyze the independent variables. The team turned to evaluating the data and sources for the independent variables used to develop the capex and opex variables in the QRA. Alexicon tested the correlation between the dependent variables and each of the two independent variables for the 726 companies in the data set. Scatterplots were created and evaluated to assess correlation, linearity, presence of outliers, etc. In the initial stages, analysis focused on cost causation and correlations. Scatterplots of the variables as well as new scatterplots based on several of our recommendations will be explained in the third section of this report and more extensively in *Appendix D*.

FINDINGS – COSTS

The effectiveness of all linear regressions depends on the proper definition of variables to provide a valid basis for establishing relationships. Accordingly, Alexicon first reviewed the definition of capex and opex costs used in the QRA to determine if a foundation had been laid for proper comparisons to be made between study areas.

1. Development of the dependent variables (capex and opex) includes arbitrary amounts and differing allocation schemes.

The goal of the QRA is to establish benchmarks for the costs to deploy and operate the high-cost loop network over which carriers provide voice and broadband services. The QRA is designed to generate cost-predictive calculations concerning:

- How much subscriber cable and wire facility loop plant and subscriber circuit equipment is needed to provide universal services in high-cost regions based on comparisons with other companies?
- How much does it cost to operate the subscriber portion of the network in comparison with other companies?

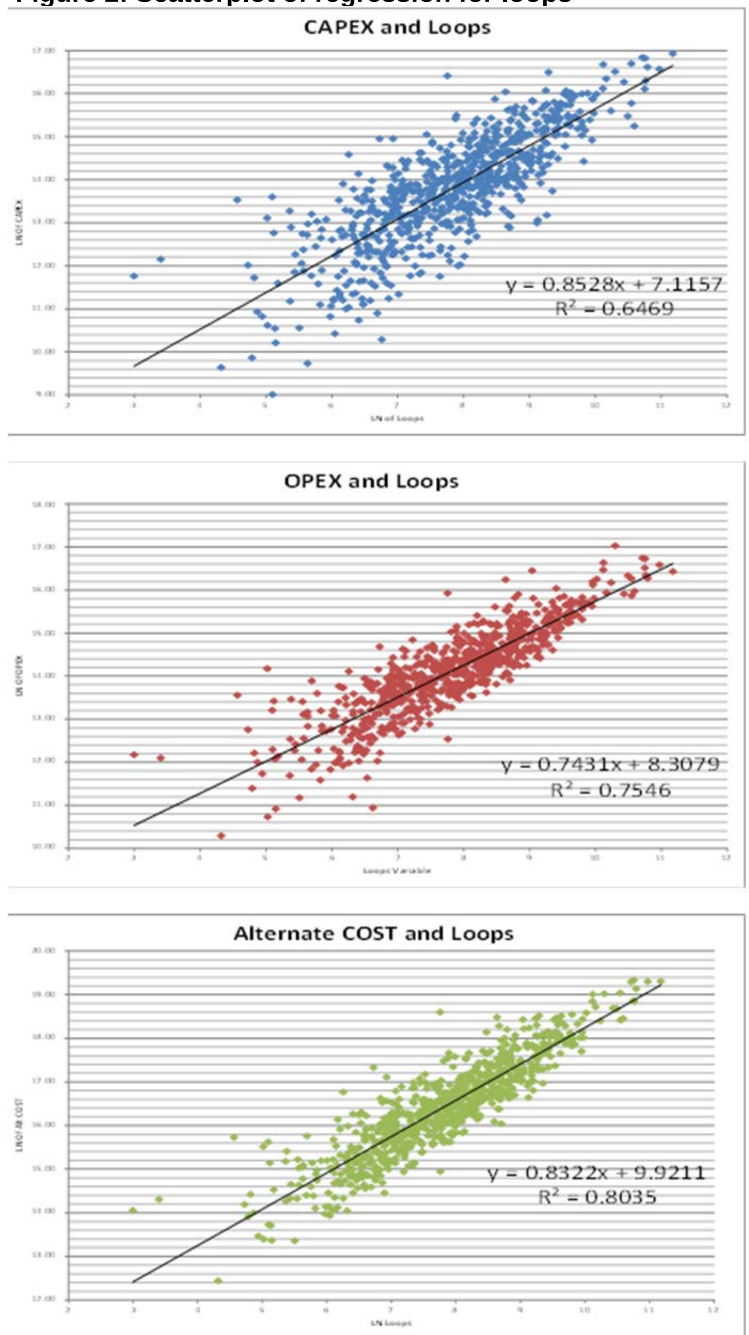
Significantly, however, instead of using the actual deployment and operating costs, the FCC used the costs from another model, namely the NECA HCL Algorithm which is used to calculate Study Area Cost per Loop. The shortcut, apparently used for convenience, introduces a host of complicating issues. The primary problem here is that the HCL algorithm costs are allocated or derivative costs. The allocation percentages vary greatly from company to company and, for the most part, are out of the companies' control since they are established by Code of Federal Regulation ("CFR") rules or state public utility commission orders. More important, the allocation percentages are not comparable, meaning that the QRA is comparing "costs" that are proportioned differently from one carrier to another.

The key takeaway is that it is impossible to establish cost relationships between a "predictor" variable and an arbitrary amount that is not the same for each company. For example, while there is a cost-causative relationship between the number of loops in a study area and the cost of cable and wire facility ("CWF") deployed in the local network; there is no cost causation between the number of loops and the depreciation rate which underlies the output variables. The number of loops cannot predict a depreciation rate arbitrarily decided by a state commission. Arbitrary, non-cost causative factors such as depreciation rates introduce an unpredictable and complicating factor that degrades the relationships between predictor variables and capital expenditure costs. See *Appendix B* for additional analysis of this problem.

2. Opex improperly includes operating taxes

The QRA seeks to identify "excessive" costs, which means expenditures that a carrier should not incur if it were responsibly operating or investing in its business. The QRA, however, embeds certain costs over which the carrier's managers have no control. The most flagrant example is operating taxes, which include property taxes, operating investment tax credits, state and federal income taxes, deferred operating taxes, and other operating taxes. Like the inclusion of depreciation rates in capex, the inclusion of widely-divergent tax rates distorts the predictive value of the QRA. And the result is not trivial as taxes can rise to nearly 40% of operating cash flows. No variable used by the FCC for scale, geography, customer dispersion, etc. can predict that the state of Nevada has no corporate income tax rate while Pennsylvania has a 9.99% corporate income tax rate.[26] Furthermore, income tax rates are significantly dependent on corporate structure as many cooperatives pay few or no income taxes while a privately-owned company may pay as much as 39%.

Figure 2: Scatterplot of regression for loops



Source: Alexicon Consulting, 2013.

Our re-creation of the QRA permitted us to analyze various issues, including the Operating Tax problem. We were able to re-run the model, combining the opex and capex dependent variables and

[26] State corporate income tax rates were obtained from the Federation of Tax Administrators at http://www.taxadmin.org/fta/rate/corp_inc.pdf

excluding Operating Taxes, to evaluate the import of an alternate regression. Despite many issues with independent variables described in the next section, the alternate quantile regression analysis returns a superior fit (0.714) compared with the FCC's opex and capex regressions. We also compared the alternate regression results on a variable-by-variable basis as shown in Figure 2. As is apparent in the Figure, the regression returns a superior fit (0.80 for our alternative model) for loops compared with the FCC's opex and capex regressions which are depicted in the top two graphics. This is an apt comparison because the number of loops is the single most reliable cost predicting variable used in the QRA. Additional graphics are found for each of the regressed variables in *Appendix D*.

The inclusion of uncontrollable costs, such as Operating Taxes, in the benchmarking process appears to be contrary to the goal of the regression analysis which is to evaluate comparable input data and then to identify where a carrier should avoid excessive expenditures.[27] With due respect, the authors submit that the FCC's justification that Operating Taxes should be included in the benchmarking process because "taxes are an expense that must be paid, just like other operational expenses"[28] is insufficient to merit inclusion in a regression analysis. The facts remain that arbitrary amounts like operating taxes cannot be predicted by the QRA and the costs themselves are out of the control of the companies. Furthermore, the inclusion of operating taxes in the benchmarking process frustrates the FCC's goal of rewarding operating "efficiency" as it distorts the QRA results. The inclusion of operating taxes in the definition of opex improperly punishes companies required to pay higher taxes while improperly rewarding companies that may not be operating as efficiently but have a lower tax burden. Operating tax amounts should not be included in the calculation of operating expense limits but rather should be included in the calculation of recoverable costs *after* any limits are imposed.

Appendix C offers more detail concerning the Operating Tax problem. In summary, of the 726 study areas, 585 qualify for high-cost loop support according to the FCC's calculations in April 2012. Our recalculation, after excluding Operating Taxes in the benchmark tool, would have resulted in significantly revised high cost loop support amounts for 80 of the 585 companies ranging from reduced support of $725,000 to restored support of $7.6 million annually. We believe that a statistical model that reduces annual support by as much as $7.6 million due to non-comparable inputs and uncontrollable costs fails to properly identify "excessive" costs.

---

[27] The FCC claims that Operating Taxes should be included in the benchmarking process because "taxes are an expense that must be paid, just like other operational expenses." See HCLS Benchmark Order, Appendix A, paragraph 79.

[28] HCLS Benchmark Order, Appendix A, paragraph 79.

FINDINGS - INDEPENDENT VARIABLES

The analysis conducted by Alexicon and EI Technologies revealed significant problems with the QRA's independent variables. The QRA uses inputs that were flawed because of (i) the use of inaccurate or outdated data; (ii) questionable or clearly flawed assumptions; (iii) weak or no cost causation which make the use of such variables problematic in a predictive model; (iv) obvious errors in the results the variables generate; (v) too few source-data points for statistical reliance; and (vi) obviously low predictive values. The following chart illustrates the problems identified for fourteen of the sixteen variables in the model.

Independent Variable	Inaccurate / Outdated Data	Flawed Assumptions	Lack of Cost Causation	Obvious Error in Result	Too Few Data Points	Very Low Prediction (R^2< .02)
Loops						
Road Miles	✓					
Road Crossings	✓					
Study Areas						
% Undepreciated Plant				✓		✓
Customer Density	✓					✓
Exchanges						
% Urban Households	✓	✓			✓	
Soil Construction Difficulty	✓		✓			✓
% Bedrock	✓	✓				✓
Climate			✓			
% Tribal Land			✓		✓	✓
% National Parkland			✓		✓	✓
Alaska			✓	✓	✓	✓
Midwest			✓			
Northeast			✓		✓	

1. Certain variables were developed with inaccurate or outdated data

If we exclude the five geographic variables—three regional, as well as tribal land and parkland—all of which are problematic in our view, we are left with eleven independent variables used as to predict capital and operating costs. Six of those remaining variables—55% of the remainder—have specific problems related to the underlying data.

Service Area Boundaries. Service area boundaries are developed using TeleAtlas Operating Company data, which does not correspond to actual company service area boundaries. Material differences between the FCC-generated service areas and actual service area boundaries have been reported by many companies. The accuracy of the variables that are Road Miles, Road Crossings, Customer Density, Percentage of Households in Urban Areas, Soil Difficulty, Percentage of Bedrock, Tribal Land percentage, and National Park percentage is dependent on the service area boundaries (see *Appendix A* for details of the variable development).

Census Block Data. Customer Density and Percentage of Households in Urban Areas are developed using census block data. In addition to the improper inclusion or exclusion of census blocks due to incorrect study area boundaries, the QRA uses a centroid-based calculation that assigns all of the census block data to a centroid (a geographic center point). This means that if a portion of the census block is within a study area but the centroid is outside of the study area, none of that census block data is included in the study area. This process can obviously distort the results. An alternative, more accurate process would be to use a clip process that would assign proportions of the census block data.

Street Mapping Data. The FCC documentation discusses the use of ArcGIS Resource Center for ESRI StreetMap. Upon further research into the dataset, the resource center documentation mentions that the data may be as old as 2007. The ESRI Street Map Premium product available from ESRI does release data every six months and is presumably more accurate than the data used.

Soil Construction Difficulty. There are two problems with the development of the Soil Construction Difficulty variable. First, the calculation uses only the "predominant" soil type. To illustrate, if an area has 51% black soil and 49% rock, the entire area is assumed to be black soil. This oversimplification of the data introduces a problematic level of inaccuracy and obviously weakens the results.

Second, Soil Construction Difficulty is based on a calculation of the dominant soil texture using a combination of databases. The thickest "horizon" within the top thirty-six inches of soil is determined, and a soil texture value for that horizon thickness is obtained from another set of data. There should be a one-to-one relationship between horizon thickness and soil texture value; distressingly, however, in the majority of cases the horizon thickness values were either blank or duplicate. A blank horizon thickness means that a texture value cannot be derived. A duplicate horizon thickness value, which occurs in some cases, results in an arbitrarily selected texture value. These troublesome issues with the source data render the Soil Construction Difficulty variable so flawed as to be meaningless, in our opinion.

2. Flawed assumptions in developing independent variables

Definition of "urban" households. The definition of "urban" used to calculate the percentage of urban households includes both census "urban areas" (50,000 or more people) and "urban clusters" (at least 2,500 and less than 50,000 people). This means that the QRA considers populations with as few as approximately 962 households to be "urban" based on the US Census Bureau average of 2.60 persons per household.[29] The inclusion of urban clusters in the definition of urban areas is questionable because there has been no support provided to demonstrate that an area with as few as 962 households is comparable with larger concentrations of population or that such an area contains the population density characteristics that result in reduced capital or operating costs. The inclusion of "urban clusters" in the definition of urban households likely skews the results and provides misleading cost relationships.

Non-linear cost relationships. The QRA is a linear regression which assumes a linear relationship between independent variables and costs. However, there is not a necessarily linear relationship between capex and certain variables, such as the percentage of bedrock. In other words, more bedrock does not necessarily mean more capex cost. The presence of bedrock may require boring or additional labor to install buried or underground cable which increases the network deployment cost. However, once the percentage of bedrock reaches a certain level, placement of buried/underground facilities becomes cost prohibitive or technically infeasible. In such cases, companies use aerial cable

---

[29] United States Census Bureau, Persons per household, 2007-2011 at http://quickfacts.census.gov/qfd/states/00000.html.

which is less expensive than buried/underground cable to deploy. This would suggest that the cost relationship between capex and bedrock, for example, is not linear and is therefore not a valid independent variable for use in the QRA. Climate may not be linear as it is generally only "extreme" conditions that create higher capex or opex costs.

Geographic regions are too broad. The Midwest and Northeast variables are defined using the US Census Bureau regional designations. A quick review of the geographic scatterplots in *Appendix D* highlights that the calculations are based on very low correlations (less than one-half of 1%) using either a 1 or a 0 to designate whether the carrier operates in the region. The regions are too broad to represent the type of homogeneous values necessary to develop valid cost relationships. A lack of homogeneity in the sample set from which a correlation was calculated can bias the correlation. Variables for the South and West were not included, presumably because of warmer climates and/or more favorable soil conditions, but the QRA does not include other geographic-related factors, such as coastal areas where the construction of plant must accommodate severe environmental conditions.

3. Variables are not cost causative

Significantly, it appears the FCC failed to establish cost-causation relationships when reviewing data correlations.[30] The major conceptual limitation of all regression techniques is that one can only try to ascertain relationships through measuring correlation. Further analysis is needed to determine an underlying causal relationship. For example, one would find a strong positive correlation between the damage that a fire does and the number of firemen involved in fighting the blaze. Do we conclude that the firemen cause the damage? No. The most likely explanation of this correlation is that the size of the fire (an external variable that is not included in the study) caused the damage as well as the involvement of a certain number of firemen (i.e., the bigger the fire, the more firemen are called to fight the blaze). Even though this example is fairly obvious, in real correlation research, alternative causal explanations should be considered. Although one cannot prove causal relations based on correlation coefficients, it is still possible to identify so-called spurious correlations; that is, correlations that are due mostly to the influences of other variables. If one "controls" for this other variable (e.g., consider only fires of a fixed size), then the correlation will either disappear or perhaps even change its sign.[31]

In another example of an even weaker causal "correlation," a University of Notre Dame Engineering professor reported to his class in 2010 on his unpublished "study" of the weather on home football weekends. He explained that his analysis showed, over approximately 60 years, there has been a remarkably high incidence of the sun shining on home football weekends compared with all other days of the week during the fall and winter. A so-called regression analysis "revealed" a greater than 80% likelihood that the sun would shine when the Irish played on campus—actually a statistic well higher than the "fit" for any of the variables used in the QRA model. It is not science, of course, but coincidence, unless one wants to assert a higher power influencing the climate for Notre Dame football games. Despite the statistical correlation, the professor's point to the class was that there is clearly no causal

---

[30] In their discussion of the Assumptions, Limitations and Practical Considerations of Linear Regressions, Hill and Lewicki state: "Multiple regression is a seductive technique: 'plug in' as many predictor variables as you can think of and usually at least a few of them will come out significant. This is because you are capitalizing on chance when simply including as many variables as you can think of as predictors of some other variable of interest."

[31] Thomas Hill and Paul Lewicki, STATISTICS: Methods and Applications (StatSoft, Inc. 2005).

relationship, that is, a home football game does not "cause" the sun to shine in South Bend, Indiana, regardless of what the computer reports.

Since the goal of the QRA is to establish cost-causative relationships in order to allow prediction of costs, it is crucial that the independent variables selected be evaluated as to whether they are true cost-causers. However, we have no evidence that the FCC ever established the cost causation of the variables. In several cases, including geographic variables, percentage of undepreciated plant, climate, and population density there is an obvious and demonstrable absence of meaningful relationships between the costs and the variables purported to predict them.[32]

Geographic variables. The FCC uses geographic variables for Tribal Lands and Parkland, as well as for three regions of the country (Alaska, Northeast, and Midwest). In the case of tribal lands and parkland, the data are based on the percentage of the service territory characterized by those designations. In the case of the other regional variables, as explained above, the FCC uses a "dummy" variable and assigns a value of 1 or a 0 to signal whether a carrier serves the region. The use of dummy regional variables and other regional identifiers are identification variables only and should be excluded from the analysis because they are not truly cost-causative. Identifying a study area as being located in what is assumed to be a higher-cost area is a "short-cut" and is not the same as establishing a cost relationship. One should identify the actual variables that cause or predict the higher costs. For example, the reasons that Alaska companies may have higher costs would include frost, rainfall, distance from major metropolitan areas, state consumer price index, population density, etc. Some of those variables are captured in other QRA inputs, so a second problem—apart from not being cost-causative—is to overweight certain inputs. In the case of the Northeast or the Midwest geographic variables, once again, it is unclear how there is sufficient uniformity among carriers across those regions in terms of cost-causation to justify inclusion in the QRA.

4. Obvious errors in results were not corrected

Percentage of Undepreciated Plant. Percentage of Undepreciated Plant appears to be falsely correlated with capex. According to the FCC, undepreciated plant is used as a proxy to capture the age of plant in service. Logically, one would expect a positive correlation between the age of telecom plant and operating expenses (older equipment would require more maintenance) and possibly little or no correlation between the age of telecom plant and capex (older plant may cost less because of inflation and the smaller capacity requirements in past years). Additionally, one could expect no correlation between age of plant and capex because the replacement cost for that plant is the same, whether the plant is 5, 10, or 50 years old. The peculiar finding is that the QRA shows no correlation between the percentage of undepreciated plant and opex (again one would expect high levels of undepreciated plant for newer infrastructure which *lowers* operating expenses), and a significant correlation with capex. In fact, the QRA's use of undepreciated plant related to capex has nothing to do with causation but reflects an improper definition of the variables. Capex, the dependent variable, consists primarily of depreciation expense (see *Appendix B* for further discussion) while undepreciated plant consists of the asset and *accumulated* depreciation. The QRA is identifying a correlation between depreciation expense and

[32] Visual inspection and R-squared values of the individual independent variables as shown in the *Appendix D* scatterplots demonstrates the lack of correlations.

accumulated depreciation which is not truly a correlation between age of plant and network costs. The false correlation requires either a change in the definition of the dependent variable (as recommended below) or the rejection of Percentage Undepreciated Plant as a proxy for the age of plant.

Alaska. The QRA results in a *negative* capex coefficient for the Alaska variable. Peculiarly, the geographic dummy variable for "Alaska" produced a capex output that maintains that costs are *46% less* to deploy local telecommunications infrastructure in that state by comparison with costs in the lower-48 states. The finding is counterintuitive considering Alaska has a frozen landmass for much of the year, extensive water-related challenges, and low population densities—all of which cause higher costs. The result is also contrary to other studies, notably a ten-year study by the United States Army Corps of Engineers which found that construction costs are 19% higher in Alaska than in the lower-48 states.[33] The Alaska coefficient is an apparent error requiring correction to assure service within some of the more challenging regions. This information was presented to the FCC in August 2012,[34] but the Bureau has not yet, to the best of our knowledge, addressed this critical and obvious error.

5. Inclusion of variables with insufficient data points for statistical analysis

In the case of certain variables, the QRA relies on relatively few significant data inputs to analyze the correlations. Illustrating this problem, null values account for the majority of the inputs for the Percentage Bedrock and Soil Construction Difficulty, which pertain to all 726 companies. When a statistical analysis relies on few data points, the outputs are less reliable and possibly not significant at all. We summarize the variables that are drawn from relatively small sets due, in the first two cases, to apparent problems in the underlying data. We cite the number of values found in the data set out of the entire 726 companies. Importantly, this is not to say that the data reported are accurate or verified, but rather that insufficient reported numbers of values are found in the data set.

- Percentage Bedrock, 203 of 726 study areas (28.0% values, 78.0% null)
- Soil Construction Difficulty, 209 of 726 study areas (28.8% values, 71.2% null)

One would expect a large number of null values for the geographic variables and that is the case.

- Percentage Tribal Lands, 190 of 726 study areas (26.2% values, 73.8% null)
- Percentage Parkland, 87 of 726 study areas (12.0% values, 88.0% null)
- Percentage Urban, 215 of 726 study areas (29.6% values, 70.1% null)
- Alaska, 17 of 726 study areas (2.3% values, 97.7% null)
- Northeast, 58 of 726 study areas (8.0% values, 92.0% null)

---

[33] *Civil Works Construction Cost Index System*, US Army Corps of Engineers, March 31, 2012. See also www.wirelessestimator.com, a wireless telecommunications industry tower construction website that shows that as of June 1, 2012 material and installation costs are respectively 24.63% and 25.18% higher than in the lower-48 states.

[34] See *In the Matter of Connect America Fund High-Cost Universal Service Support*; WC Docket Nos. 10-90, 05-337; Petition for Waiver of Matanuska Telephone Company; filed August 28, 2012.

6. Inclusion of variables with very low predictive value

Seven of the sixteen independent variables used in the QRA have very low predictive value. The coefficient of determination, or R^2, is a commonly-used statistic to evaluate the model "fit." The R^2 value indicates how much of the change in the dependent variable is explained by changes in the independent variable(s). For example, if one were to run a linear regression to determine to what extent a change in the number of loops correlates to the change in the capex. The analysis results in an R^2 value of 0.52, which means that 52% of the change in capex is explained by the change in number of loops. In the QRA, seven of the independent variables returned calculated "fits," or R^2 values, of less than 0.02. In other words, those variables each explain less than 2% of the change in capex or opex. The revised overall QRA reports a pseudo-R-squared of 62% for the operating costs per line and 67% for the capital expenditures per line, which means that its overall predictive value remains low.

Other General Findings

The procedures and data that are excluded from any analysis are important, and possibly more so than those that are included. The following section summarizes apparent major omissions in the QRA.

1. Omission of Important Cost Causation Realities

The challenge in creating a model is to understand and capture real-world factors. The QRA omits factors that are apparently important in explaining a carrier's costs in any given year.

Amount of broadband deployment. The extent of broadband-capable network is one of the most important factors driving the cost of network deployment. Many carriers are incurring significant expenses in this network evolution. The model does not capture the cost differences associated with running a broadband-capable network compared with an older, copper legacy network. Nor does the model capture the entirely appropriate costs which are elevated in a transition from an older to a newer network, that is, when duplicative systems must be operated and maintained. This real-world challenge results in very different capital and operating expenses compared with the relatively low-cost operation of a legacy network. In light of the FCC's mandate to focus Universal Service Funds on broadband service, the absence of any measure of broadband deployment as a cost variable of the QRA is difficult to understand.

Percentage of underground vs. aerial facilities. It is significantly more expensive initially to bury cable than to deploy aerial plant. However, good reasons exist to trench outside plant, including zoning regulations, weather conditions, and the reduction of long-term operating costs. The network architecture decision to deploy underground or aerial cable has significant impact on the initial capital expenditures and future operating costs of a carrier. No such variable exists in the model.

Weather. The "climate" variable used in the QRA is only a frost index. There are other significant factors such as high winds, water, flooding, etc. that determine not only the relative necessity of burying cable but the expense to do so. If a carrier operates in a region where severe weather is common, buried plant may be demonstrably the best customer solution in assuring safety and other welfare, as well as the avoidance of repeated rehabilitation of damaged plant in the wake of extreme weather.

Lumpy capex investment cycles. Most carriers have lumpy investment cycles, as costs are well higher in certain years than in others. The reason of course is that it is more efficient to perform significant outside plant construction in a single year or several years. Like Verizon or AT&T that dedicated extraordinarily high levels of capex in a concentrated period of time to upgrade loop plant, small carriers do the same. The model fails to capture such an effect and could penalize carriers for efficient concentrations of upgrades.

Topography and other construction difficulty factors. Despite the noted problems with the FCC's development of a soil construction difficulty and percentage bedrock variables, the authors believe that construction difficulty factors such as soil type, topography, water, etc. are cost-causative factors that should be included in a cost model, assuming more reliable data sources. Another significant cost-causative variable relates to delivery of equipment and technical personnel for installation or maintenance in remote locations which create incrementally higher expenditures; these include Alaska or very rural regions of some states as well as island service territories in the continental U.S. or Hawaii.

2. Lack of comparability

It is important to understand from the outset that the FCC's commentary that the QRA provided a comparison of costs with "similarly-situated companies" does not mean that carriers of certain sizes are compared nor does it mean that companies in Alaska are compared with each other.[35] It means that the carrier is compared with all the other 725 rate-of-return carriers. This loose application of the principle related to comparisons with "similarly-situated peers" appears contrary to both the FCC's stated objectives with respect to the model *and* the FCC's explicit delegation/instruction to the Bureau.[36] Comparability of data is a serious issue in assessing the costs of rural telecommunications carriers. The 726 study areas involved in the QRA range from single-exchange carriers with less than 200 access lines to carriers that are over 100 times that size situated in vastly differing geographic areas. There are two main ways to ensure comparability of data. The first is to capture all or the vast majority of the causes of differences between costs in the study areas. The second is to segregate the data into more comparable subsets (i.e., compare study areas of similar size, geography, broadband deployment, etc.). The QRA does not attempt to do either. Also, as explained above, there are major cost differences between carriers that make them affirmatively not comparable, including companies paying little or no operating taxes and others that pay as much as 39% or companies whose depreciation rates are set at sharply different levels from those of other carriers.

3. Outliers

The authors also evaluated outliers, which are atypical, infrequent observations. Because of the way in which the regression line is determined, outliers have a profound influence on the slope of the

---

[35] USF/ICC Transformation Order, Appendix H, footnote 1: "The term 'similarly-situated peers' means that, based on data from all the carriers in the analysis, if there were (hypothetically) 100 study areas with independent variable values that were nearly the same as those with the study area in question, 90 of them would be expected to have values equal to or less than the 90th percentile prediction. It does not mean the carriers with the most similar number of loops (or values of the other variables)."

[36] USF/ICC Transformation Order, paragraphs 210-226.

regression line and consequently on the value of the correlation coefficient. Although many assumptions of multiple regression analyses cannot be tested explicitly, gross violations can be detected and should be addressed either in the formula or through exclusion from the analysis. The reason, of course, is that outliers can seriously bias the results by "pulling" or "pushing" the regression line in a particular direction, thereby leading to insufficient regression coefficients. The FCC states that the QRA "solves" the problem of outliers, but the reality is that quantile analyses mitigate or prove more "robust" in addressing the problem rather than "solving" it.[37] The FCC declined to remove outliers due to concerns about the judgments required to define outliers. While this concern is understandable, it ignores the fact that quantitative approaches to outlier removal are common practice. The authors were not able to generate a significantly higher correlation in removing outliers, but note that the so-called "father of the quantile regression" commented in a submission to the FCC about the QRA regarding the potential "outliers . . . [that] can be highly influential in [linear and quantile] forms of regression and can do serious damage to the fitted model."[38] The lack of any type of examination of the effect of outliers on the QRA results represents a potentially serious procedural flaw.

---

[37] See USF/ICC Transformation Order, Appendix H, para. 7. The reality is that quantile regression works to arrive at a median result based on the sum of the absolute residuals rather than squares of the residuals; so like a median value rather than an average value, the very high or very low extremes affect the results to a lesser degree. See, e.g., *Outliers,* summary compilation of statistical sources, available at http://www.nd.edu/~rwilliam/stats2/l24.pdf, p. 18; John Fox, Regression Diagnostics, McMaster University, 2009; George Bohrnstedt and David Knoke, Statistics for Social Data Analysis, Peacock Publishers,1982.

[38] National Exchange Carrier Association et al. Comments, WC Docket No. 10-90 et al., at App. E, 8 (filed Jan. 18, 2012) (Roger Koenker, "Assessment of Quantile Regression Methods for Estimation of Reimbursable Cost Limits") (Rural Association Comments).

Section 2: Perspectives on the QRA's role in meeting policy obligations

A study of the FCC's QRA might lose the non-technical observer in the weeds. Statisticians and economists debate about assumptions, coefficients, confidence factors, and underlying data. We believe that the previous section highlights a relatively straightforward set of problems. However, there is a larger question that requires a policymaker to step back.

The question is *whether the QRA is a good tool, as employed, in realizing the statutory policy goal of "specific, predictable and sufficient Federal and State mechanisms to preserve and advance universal service."* This is not a question about statistics or databases, although the previous section of this White Paper raises important questions about the sufficiency of the model. Nor is it a question that should ultimately be answered by technical people or cost consultant experts. This is a more fundamental question that is for policymakers who realize that there is an affirmative and important goal in universal service. The policymaker question is whether the QRA, as a tool, advances the goal of realizing universal service, or, in slightly different words, whether the benefit of the QRA is worth the risk and the cost associated with the statistical mechanism?

We suggest that there are several issues in answering this fundamental question.

- *Benefits of the QRA.* It is important to understand the policy-purpose the QRA is designed to accomplish, which moves the FCC to implement such a mechanism.
- *Risks arising from use of the QRA.* The risks and costs associated with the QRA should be carefully examined.
- *Broader assessment of the affirmative goal regarding USF.* The statutory goal has not only been spelled out in the Telecommunications Act of 1996, but has been interpreted by the FCC previously. It is helpful to review the previous assessments to gain perspective on the current reforms.

Benefits of the QRA

The FCC clearly articulated at the beginning of its USF/ICC Transformation Order the four "principles" employed in reforming USF. It appears that there is one affirmative goal, which is to "modernize USF and [intercarrier compensation] for Broadband" in such a way that "[u]nserved communities across the nation cannot continue to be left behind."[39] The remaining three principles relate

[39] USF/ICC Transformation Order, para. 11: "▪ Modernize USF and ICC for Broadband. Modernize and refocus USF and ICC to make affordable broadband available to all Americans and accelerate the transition from circuit switched to IP networks, with voice ultimately one of many applications running over fixed and mobile broadband networks. Unserved communities across the nation cannot continue to be left behind.
▪ Fiscal Responsibility. Control the size of USF as it transitions to support broadband, including by reducing waste and inefficiency. We recognize that American consumers and businesses ultimately pay for USF, and that if it grows too large this contribution burden may undermine the benefits of the program by discouraging adoption of communications services.
▪ Accountability. Require accountability from companies receiving support to ensure that public investments are used wisely to deliver intended results. Government must also be accountable for the administration of USF, including through clear goals and performance metrics for the program.
▪ Incentive-Based Policies. Transition to incentive-based policies that encourage technologies and services that maximize the value of scarce program resources and the benefits to all consumers. We have also sought to phase in reform with measured but certain transitions, so companies affected by reform have time to adapt to changing circumstances.

to assuring that the monies assigned to the program are responsibly and accountably utilized—fiscal responsibility, accountability, and incentive-based policies.

The FCC employed the QRA to reduce certain funding levels.[40] However, the reductions for the high cost loop fund or HCL are small, relative to the overall size of high-cost USF. According to the USF/ICC Transformation Order's Appendix B, *Quantile Regression Cost per Loop*, released on April 25, 2012, the annualized disallowance of expenses is approximately $44 million in capex funding and $50 million in opex funding. The total disallowance is approximately $94 million or 2.1% of the total high cost fund and 4.7% of the funds dedicated to rate-of-return carriers.[41]

RISKS ARISING FROM THE USE OF THE QRA

Among the most important risks is the increased uncertainty for small carriers that have limited access to capital and relatively low visibility about ongoing cash flows as terminating access revenues (ICC) decline to zero when "bill and keep" is fully in effect in 2020. The QRA raises risks that even relatively small changes will result in underfunding for networks serving customers when a carrier is small and financially vulnerable. Policymakers should weigh the benefits of saving and reallocating 2% of the overall fund or 4.7% of the rate-of-return funding against the increased potential for harm if the QRA results in a miscalculation affecting customers of small carriers in costly service regions.

Risk that the model is inaccurate

The first risk is that an unproven QRA "model" is used where certain of the input data are wrong, and this appears to be the case at the present. There are meaningful data and design flaws in the current model as explained in detail in the initial section and in the appendices at the end of this report.

Risk arising from automatic disallowance of costs

Using a model automatically to disallow some cost recovery for certain carriers is likely to create other problems. This is not to say that a regression model is useless, but that it appears unwise to rely on a QRA to generate results that are sufficiently accurate such that recovery can be disallowed automatically. The varying carrier sizes, diversity of the operating environments, and the financial vulnerability of relatively small carriers in primarily high-cost regions, all make it difficult to model appropriate funding.

This same risk was studied at the time of the initial major reform of universal service, which was implemented by the FCC in 2001. To prepare for that earlier reform, the FCC relied upon the Federal-State Joint Board on Universal Service, which then assigned the analytical work to an interdisciplinary group known as the Rural Task Force ("RTF").[42] The RTF membership was broad, diverse and highly

[40] In the FCC's April 25, 2012 Order, the regressions were modified to generate estimated capex and opex costs. The new approach replaced the original plan to use eleven different regressions. *See* April 25 Regression Order, para. 12.

[41] The capex disallowance is estimated to be $43,982,450; the opex disallowance is estimated at $49,968,397; and the total disallowance is estimated to be $93,950,852. The totals are summations of the columns found in the table in Appendix B, Quantile Regression Cost per Loop (CPL) of the HCLS Benchmark Order.

[42] Federal State Joint Board on Universal Service, Report and Order, CC Docket No. 96 45, FCC 97 157, 12 FCC Record 8776, (1997) (May 8 Order), Paragraph 253. Also see Texas Office of Public Utility Counsel v. FCC, No. 97060421, 1997 WL 556461 (5th Cir. July 30, 1999).

skilled as outlined above in footnote 9. The RTF generated six White Papers, beginning in September 1999; one of the most cited reports was released in January 2000 as the RTF White Paper 2, which was a 66-page document regarding the special characteristics of rural telephone service. That document provided a detailed analysis of the differences between rural and non-rural carriers, and importantly *among* rural carriers.[43] In September 2000, RTF provided its recommendation to the Joint Board. The RTF summarized its findings in that final document:

> "Congress, the FCC and the Joint Board have each concluded that universal service mechanisms and policies applying such mechanisms must be flexible in recognition that market and operational factors associated with Rural Carriers may be substantially different from those associated with non-Rural Carriers. For the most part, however, the precise scope and magnitude of those differences had not been documented. Recognizing this gap in the evidentiary record, the Task Force undertook a detailed study of the 'rural differences.' Conclusions from that study are summarized in White Paper 2, 'The Rural Difference,' released by the Task Force in January 2000. White Paper 2 analyzes publicly available national data assembled for the first time, to systematically compare and contrast Rural Carriers and non-Rural Carriers. *Equally important, the analysis conducted by the Task Force documents a substantial diversity among Rural Carriers themselves. An understanding of the differences between Rural Carriers and non-Rural Carriers and the diversity among Rural Carriers is key to designing appropriate mechanisms and policies which will allow the fulfillment of the 1996 Act's universal service principles*."[44] (Emphasis added.)

As a result of the differences among rural carriers, the RTF recommended the rejection of a forward-looking model and urged the use of a modified embedded cost model.[45] Importantly, the QRA is not a forward-looking model, but the QRA is a mechanism that assumes similarities among the operating and capital characteristics of the rural carriers. This approach is in sharp contrast with that of the RTF, which explicitly rejected the assumption that there were significant similarities among rural carriers, arriving at the conclusion that there was "substantial diversity among Rural Carriers," all on the basis of a fact-driven study. Those RTF findings were a key element of the analysis that led to the 2001 USF reform, which used the individual carrier's actual costs for defining USF support.

The FCC today is using the QRA to define appropriate carrier-specific costs. The argument is that support will be "more predictable for most carriers because the new rule discourages companies from exhausting the fund by over-spending relative to their peers."[46] However, to the best of our knowledge, no new study or new data have surfaced in the last twelve years to overturn the findings of the RTF that there are substantial differences among rural carriers with respect to their costs. If that is the case, then there continues to be a high likelihood that there are significant differences among rural carriers that the use of a model fails to capture—at least as a definitive tool automatically in disallowing costs.

---

[43] Rural Task Force, White Paper #2: The Rural Difference, January 2000, available at http://www.wutc.wa.gov/rtf/old/RTFPub_Backup20051020.nsf, (hereafter, "RTF White Paper 2"), p. 5.
[44] Rural Task Force Recommendation to the Federal-State Joint Board On Universal Service, September 29, 2000, available at http://www.wutc.wa.gov/rtf/old/RTFPub_Backup20051020.nsf, (hereafter, "Recommendation of the RTF"), pp. 10-11.
[45] Recommendation of the RTF, p. 4.
[46] HCLS Benchmark Order, para. 41.

Simply stated, there is a reasonable and significant risk that the QRA model is wrong in calculating an individual carrier's appropriate costs. The FCC appears to believe that the model can work based on an assumption that is contrary to the fact-based findings of a major task force of experts that reviewed this matter just over a decade ago.[47] In the absence of any proof to the contrary, the RTF's findings highlight the risk that a carrier could receive insufficient funding in its specific service region as it attempts to serve its rural customer base.

Risk that the QRA creates more uncertainty and unpredictability

The statutory goal is to provide USF in such a way that investment and services can occur in regions which are unserved or underserved. The FCC has assured that it seeks to provide "predictable funding" for eligible rural carriers.[48]

The FCC has stated that the QRA is similar to the cap on the High Cost Loop fund. That is, a rural carrier cannot know year-to-year whether the National Average Cost Per Loop ("NACPL") would result in a disallowance of some funding to the carrier. The logic is that the QRA is unpredictable in a way that is similar to the NACPL. The truth, however, is that previously a carrier could reasonably estimate its own costs and the NACPL, which was rising each year at relatively predictable levels—about $21 each year. By contrast, the QRA is a complicated statistical analysis that cannot be predicted—at least at the present—with any degree of confidence. And, from 2014, it will be recalculated annually. The variables and the predictability of the QRA are materially more complex than the NACPL estimation process.

As a result of the far more complicated analysis of the QRA, the effect of the use of the model as an automatic disallowance is to create a much higher degree of unpredictability and to incent very conservative levels of spending by an individual carrier so that it does not risk shortfalls in recovery on its high-cost spending. Then, if most carriers take this approach each year as would be rational, each subsequent year becomes more conservative and there is a potential "race to the bottom."

Some policymakers may believe that cost-centric discipline is precisely what the reforms are seeking, but this specific implementation process is potentially in conflict with statutory policy. In this new QRA-influenced world, carriers are not motivated to invest what is sufficient, but to focus on avoiding the more significant unpredictability in the capital and operating environment. It is in fact unpredictability—arising from a complex and uncertain statistical tool—that could, or arguably will, prompt *underinvesting behaviors in costly unserved or underserved areas*. The symptoms of potentially sub-par investments are explained below.

---

[47] *See* HCLS Benchmark Order, para. 20: "The independent variables in this study are those that *we believe* correlate with each carrier's costs, are currently available to the Commission, and exist for all study areas in the regression analysis" (emphasis added); footnote 40: "We understand that carriers serve business as well as homes, but we do not have business information with the same urban breakout as housing units. *We are comfortable with the assumption* that businesses and homes are similarly distributed throughout study areas for rate-of-return carriers" (Emphasis added.)

[48] USF/ICC Transformation Order, para. 207.

If the QRA is not tested regularly, including by thoughtful reviews, then damaging policy outcomes are to be expected. Carriers will invest only if there is a rational and predictable opportunity for recovery of their invested capital, and they will err on the side of conservatism. The QRA is problematic precisely because it injects far more uncertainty into the investing environment, and makes it possible that a carrier could lose a substantial amount of support because of unknowable statistical calculations . . . or mistakes.

BROADER ASSESSMENT OF THE AFFIRMATIVE GOAL REGARDING USF

The FCC has stated the affirmative goal that USF should be modernized so that broadband can be more universally available, particularly in unserved or underserved regions. Broadband investment requires substantial changes to the plant between central offices and homes or businesses.

The fact that the FCC is proposing automatic *reductions* to the funding of some carriers' high-cost loops can only result in increased caution, greater effort to focus on more profitable customers, and ultimately diminished investment in unserved or underserved regions. The result is predictable and is already beginning to unfold. The authors of this White Paper have many clients who are already refocusing their businesses by reducing capital commitments in high-cost areas as a result of the uncertainty significantly arising from the use of the QRA. In a more systematic survey, the National Telecommunications Cooperative Association ("NTCA") conducted a 2013 inquiry among its membership, which are small rural telecommunications companies (about half of which are cooperatives); the survey found that 69% of the respondent carriers were postponing or cancelling "fixed network upgrades as a result of the uncertainty surrounding [the USF/ICC Transformation Order]."[49] Confirming the NTCA findings, the two major lenders to rural carriers, CoBank and the Rural Utilities Service, report sharply lower lending for network infrastructure over the last year. CoBank reports that it is making no new infrastructure loans in light of the challenging and uncertain investment environment in the wake of the FCC's recent reforms.[50] The other lender is the U.S. Department of Agriculture's Rural Utilities Service ("RUS") which has been able to loan its entire capacity of available funds virtually every year until 2012, when it was able to lend only 11.6% of the $690 million that was available.[51] Further, of

---

[49] National Telecommunications Cooperative Association, "Survey: FCC USF/ICC Impacts: Summary of Results," February 2013, available at www.ntca.org.

[50] January 23, 2013, conversation between Michael J. Balhoff and Robert F. West, CoBank, Senior Vice President, Division Manager; see, also, Letter of Robert F. West to FCC, Marlene H. Dortch, May 18, 2012, available at https://prodnet.www.neca.org/publicationsdocs/wwpdf/0511cobank.pdf. "CoBank is concerned about the negative impact the USF/ICC Transformation Order (the Order) is having on investment in rural broadband. The various caps and limitations on universal service funding and inter-carrier compensation, especially for rate-of-return carriers, are making it increasingly difficult for us to extend credit for the purpose of deploying ubiquitous rural broadband networks. . . . It is a stated objective of the Commission to support the deployment of rural broadband. Unfortunately, we view many of the provisions of the Order, especially the use of QRA, as antithetical to that goal. Affordable broadband for all Americans cannot be achieved without increasing the funding spent to support broadband deployment. The rate-of-return regulated Rural Local Exchange Carrier has historically done the lion's share of the work in deploying truly robust broadband in rural America. Instead of trying to find ways to cut and curtail support to these carriers, we continue to believe the Commission's goals would be better served in finding ways to help these carriers continue to succeed in their decades-long mission of bringing modern telecommunications services to their subscribers."

[51] The RUS agency staff reports that it has fully obligated the appropriated funds 2008-2011. In 2007, the RUS obligated 81.3%. The staff records indicate that the RUS was at or near full obligation in the several years preceding 2007. The drop-off in 2007 was the end of the Rapid Refinance program (2004-2006) when the RUS was refinancing its loans.

another $736 million available for RUS broadband loans, only 9.4% ($68.9 million) was drawn down in 2012.[52] The startling message appears to be that investors—corporations, or debt or equity investors—are not likely to increase capital investments in a time of sharply lower support and uncertain public policy.

The question for policymakers to ask is whether they derive sufficient benefits from "saving" or "reallocating" 2% of the fund each year. Second, they should ask whether the carriers are made more "accountable" for achieving policy goals or whether the carriers are incented to behaviors that are harmful to those goals. Parenthetically, they might ask whether this mechanism relieves the regulator of its important accountability role in assuring and testing the realization of universal service goals. Finally, policymakers should ask whether the QRA creates more predictability or less in the investment process.

The findings of this White Paper suggest that the QRA fails to advance the universal service goals of the United States.

[52] The United States Department of Agriculture / Rural Development, "The Telecommunications Program," presentation by RUS Deputy Administrator Jessica Zufolo to the National Association of Regulatory Utility Commissioners, Washington, DC, February 2, 2013, slide 5. See, also, "Vilsack, RUS Meet With Genachowski To Discuss The Need For More Changes In Implementation Of USF-ICC Transformation Order: Warn Of Unintended Consequences And Need For USF-ICC Support To Be Sufficient and Predictable," Independent Telecom Report, Volume 12, Issue 3 (February 18, 2013), pp. 3-5); "In the meeting [with FCC Chairman Julius Genachowski and his staff], [Secretary Vilsack and] USDA officials noted that demands for RUS loans dropped dramatically in 2012. RUS reported "demand" for only 37 percent of the funds that were actually appropriated by Congress. USDA cited the reductions in USF and ICC that will result from the implementation of the FCC's Transformation Order as the reason for the decline in loan applications. Rural carrier advocates have noted that the reduced loan activity reflects the adverse impact of the FCC Order on infrastructure investment and rural community economic development." The figures were also reported in an ex parte filed at the FCC on February 15, 2013. The reconciliation is that the "demand" for loans was reported as 37% according to Secretary Vilsack, but the RUS actually "obligated" the amounts reported by Ms. Zufolo.

Section 3: Recommendations

Suggested Improvements to the QRA

This section proposes constructive improvements to the QRA. In addition, we recommend that an improved QRA should be used in a way that is different from what has been ordered by the FCC. Specifically, a predictive model could identify carrier costs that appear excessive so that "accountability" can be improved in funding universal service. But the model should be used to highlight *potential* outliers rather than automatically presume that the statistical tool *proves* excessive spending in the case of the carriers in the top 10% of the model's output.

This third section's purpose is to focus on how to improve the QRA in whatever way it is used, but it is our view that the data and model will never be sufficiently accurate to be relied upon as the ultimate arbiter of funding.

Recommendations about the Use of the QRA

This study identifies specific flaws in the current version of the QRA. Further, this study proposes constructive approaches by which the QRA can be improved.

There are three other larger problems with the *automatic disallowance of funding* for certain carriers. Those problems are, first, that any model, within so disparate an industry, will likely fall short—as an automatic disallowance—in predicting an individual carrier's costs. Models are always tools and will presumably fall short of reality. More to the point, the RTF in 2000 was clear from its study that there were significant differences *among* the rural carriers, such that models were problematic. Additionally, the industry is today undergoing a transition from copper and circuit-switched investments to fiber and IP architectures, which requires a change in the outside plant. The transition means that comparisons of costs between carriers with legacy operations and carriers with new investment in higher bandwidth services are problematic, especially when the higher-cost investment in broadband is the stated policy goal.

The second reason to avoid automatic disallowances is simpler and more direct. The QRA is untested. To use such a mechanism without the benefit of experience is questionable policy, particularly when one considers the characteristics of service in high-cost regions. If the QRA is in fact a good predictive mechanism—and there is again reason to believe that no model will be sufficient without complementary human judgment or some other corroboration—it will be proven over time through regular reviews.

The third reason that automatic disallowances are not a sound approach is that it unnecessarily puts the statutory mandate at risk. It can reasonably be argued that the model is not "specific" since it assumes that it is correctly predicting costs on the basis of approximate inputs from 725 other carriers, and does so without significant testing or the opportunity for audit or review. The model is most assuredly less "predictable" since carriers previously were able to calculate relatively simple variables—their own costs and the NACPL that was increasing by approximately $21 annually; the QRA is decidedly unpredictable because no carrier today can estimate what will be allowed or disallowed in an annual recalculation of a multi-variable equation used for 726 carriers. Within the last weeks, the FCC released a new calculation of the QRA which raised the number of carriers above the capped levels by about 60% compared with the benchmark model from April 2012. Finally, it can be argued that the QRA will eliminate "sufficient" investment, particularly because it incents behaviors in which carriers will be increasingly cautious each subsequent year; and rational carriers industry-wide will focus on an unpredictable recovery mechanism rather than on the affirmative policy related to sufficient and defensible broadband investment.

In our view, these problems arise, not from the use of a good or bad or a mediocre QRA, but from the automatic disallowances triggered by the QRA. Again, the QRA is "saving" approximately 2% of the entire high-cost fund. In weighing the benefits against the risks, it is our proposal that the QRA should be improved and then used as a tool to trigger reviews of apparently high costs incurred by carriers.

SUMMARY OF RECOMMENDATIONS

- *QRA should be modified to account for demonstrable weaknesses*
  - The opex and capex calculations should be combined into a single output to avoid allocation problems associated with the current model
  - Improved data sources should be identified and used
  - Costs that cannot be controlled by the carrier, such as taxes, should not be employed in computing appropriate operating and/or capital expenses
  - Depreciation expense, because of the different schedules used by carriers, should not be used as a proxy for capital investment
  - Geographic variables, for Alaska, the Northeast and Midwest as well as for tribal lands and park lands, should be eliminated, and, if necessary, other truly cost-causative variables added
  - Other variables should be added to accommodate percentage of the plant that is broadband capable, transition periods from legacy to broadband, severe weather factors that prompt a carrier to bury plant or possibly elevate plant to ensure that electronics are above flood levels, and extraordinary delivery and installation costs in remote locations
- *The QRA should be used as a tool to trigger evaluation of costs that might be high rather than as an automatic disallowance of certain costs*
  - If the QRA is reviewed in the process of evaluating certain carriers' costs, the model will provide an ongoing and better source of information for regulators and for companies about the specific implementation of universal service and the sufficiency of the program
  - A review process will give carriers greater confidence that they can invest and will not be subjected to an unpredictable and uncertain investment cycle; a review process will allow the industry to rely on a predictable process and set of mechanisms in support of policy-driven investment
  - The review might be conducted by the FCC or at the state commission level or possibly by some other entity such as the Universal Service Administrative Company[53]

---

[53] This proposal is consistent with the proposed NARUC Draft Federalism Principles prepared by theFederalism Task Force of NARUC, February 15, 2013, available at http://www.naruc.org/committees.cfm?c=69; "States and state regulators are closest to their citizens and are thus best positioned to determine the level of protection and

- *An interdisciplinary committee should be established to advise the state commissions or the FCC or USAC, as appropriate*
  - The committee might be chosen to represent various parts of the industry, including large carriers, competitors, wireless operators, rural carriers, professional organizations such as the National Exchange Carrier Association; the committee should have expertise in understanding the financial factors affecting the industry, as well as economics and statistics
  - The committee should review at least on a summary level, the "triggered reviews" over the last year so that a detailed understanding of the progress of USF can be achieved

---

service that should be available to the users of communications services regardless of the technology used to provide those services. For that reason, the states should take the lead in the key areas enumerated below (i.e., the principles of cooperative federalism). . . . 5. Universal service: States and state regulators are best positioned to determine the optimum method for ensuring the availability of broadband and voice services. 6. Regulatory diversity: Regulation should be "functional" and decided after reviewing and evaluating constitutional and statutory state and federal roles and exploring multiple points of view. The regulator closest to the end user should speak for that user; Federal and state regulators should seek multiple points of view on issues, including using the Joint Boards to ensure that state and end user needs are heard and understood. 7. Evidence-based decision making: regulatory decisions should be made only after open discussion, with all parties listening to all others. 8. Broadband access, affordability, and adoption: *state regulators should take the lead role in determining state requirements for broadband deployment and consumer protection. States and state regulators are best positioned to determine where broadband is needed and how it should best be deployed, as well as to evaluate competitive choice.* States and state regulators should assume a lead role in defining and ensuring consumer protections for broadband service, including privacy issues, fair and accurate billing, open access, and complaint review. States and state regulators are in the best position to track and evaluate service reliability and should have access to outage reports, service restoration information, and the other data necessary to ensure network reliability and consumer safety." (Emphasis added.)

Appendix A: Analytical Process

This appendix provides guidance on how Commission staff calculated the variables used in its HCLS benchmark calculations. These steps reflect our approach in rebuilding the FCC's QRA model.

1. This *Appendix A* first documents the process used to develop the variables requiring geoprocessing steps, and then describes how Commission staff calculated the other variables.
   - For the geospatial parameters, each process can be manually performed in ESRI ArcMap, or can be implemented as an automated process. This document will indicate where Commission staff used an automated process. Commission staff used the Python programming language for its scripts, but other scripting programming languages may be used.
   - The original geospatial data sets may have to be preprocessed in order to be used in some of the following processes. For instance, some data sets were too large for a specific ArcGIS tool, so a subset of the data sets was retrieved first before using the ArcGIS tool. In addition, some data sets are not available as a single national file, such as the National Hydrology Database ("NHD") used to calculate stream crossings. In these situations, Commission staff used a subset of the NHD data set based on the corresponding location of the study area.

2. The following ArcGIS 10.0 geoprocessing tools were used throughout the process and are referenced in this workflow. The links provide a full description of the tool's functionality and use.
   2.1 Feature to Point: http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#//00170000003m000000
   2.2 Clip: http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/Clip/000800000004000000/
   2.3 Feature Vertices to Points: http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#//00170000003p000000
   2.4 Intersect: http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/How_Intersect_works/00080000000z000000/
   2.5 Identity: http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/Identity/00080000000n000000/
   2.6 Frequency: http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/Frequency/00080000001w000000/
   2.7 Summary Statistics: http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/Summary_Statistics/00080000001z000000/
   2.8 Symmetrical Difference:http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#//00080000000r000000
   2.9 Project: http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/Project/00170000007m000000/
   2.10 Near: http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#//00080000001q000000
   The TeleAtlas data do not contain boundaries on a study area basis, but instead contain boundaries for carriers on an operating company number ("OCN") basis. In order to

match the OCN polygons to study area boundaries, Commission staff intersected the OCN polygons with the state boundaries to create OcnSt polygons, which were used to create certain values (described below), which were then summed up to the study areas using the OcnSt-Study area cross reference published at http://transition.fcc.gov/wcb/iatd/neca.html. Commission staff created the OcnSt polygons because there are some exchanges that straddle state lines, and this allowed Commission staff to associate the pieces with the proper study area later in the process. In the majority of cases, there appears to be a one-to-one match between OCN polygons (i.e., boundaries) in the TeleAtlas data set and study area boundaries, but in some cases there is not. Occasionally, a single OcnSt polygon had to be split up into multiple study areas. In these circumstances there were non-rate of return study areas that Commission staff manually excluded from the OcnSt polygons. Also, there were some cases where multiple OCN polygons had to be added together to form a whole study area.

3. Data processing flow for each OcnSt polygon
   3.1. Determining the WGS84 UTM Zone of a OcnSt polygon. The UTM Zone of an OcnSt was calculated based on the OcnSt polygon and used to generate a projected coordinate system in order to calculate geometries (area and length) of features. When clipping other features to an OcnSt polygon, the resulting data set will be in the UTM projection. The UTM zone was described at http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/What_are_projected_coordinate_systems/003r0000000p000000/. Commission staff calculated the UTM Zone as described below.
      3.1.1. Commission staff wrote a Python script to execute the following geoprocessing steps
      3.1.2. Determine the geographic boundary of the OcnSt polygon (map extent) http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#//00s900000011000000.htm
      "top": the max latitude of the boundaries of the OcnSt polygon
      "bottom": the min latitude of the boundaries of the OcnSt polygon
      "left": the max longitude of the boundaries of the OcnSt polygon
      "right": the min longitude of the boundaries of the OcnSt polygon
      3.1.3. Determine the center point of the geographic boundary
      clat : the Latitude of the center point of the OcnSt polygon
      clon: the Longitude of the center point of the OcnSt polygon
      clat = (top - bottom)/2.0 + bottom
      clon = (left - right)/2.0 + right
      3.1.4. Determine the UTM Zone of the OcnSt polygon
      int(x): the whole number of the "x"
      floor(x): the largest previous integer of the "x"
      the UTM Zone = int(floor(((clon+180) - floor((clon+180)/360)*360-180 + 180)/6)) + 1
   3.2. Calculating Total Area
   Total Area is the total area of the OcnSt polygons in the study area. The calculation process used by Commission staff follows.
      3.2.1. Commission staff wrote a Python script to execute the following geoprocessing steps
      3.2.2. Used the ArcGIS Clip tool to clip urban areas to the OcnSt polygon
      3.2.3. Used the ArcGIS function Calculate Geometry to calculate area (Total Area) of the OcnSt polygon in square miles

3.2.4. Summed the total area values using the OcnSt-Study area cross reference published at http://transition.fcc.gov/wcb/iatd/neca.html.

3.3. Calculating Water area, Land area and Pctwater

Water area is the area in the study area that consists of water bodies as defined by the US Census Bureau. Land area is the study area's total area minus water area. (Total area is calculated in 3.2.) Pctwater is water area divided by total area. The documentation about water area can be found at http://www.census.gov/geo/www/GARM/Ch15GARM.pdf The calculation process used by Commission staff follows.

3.3.1. Commission staff wrote a Python script to execute the following geoprocessing steps
3.3.2. Used the ArcGIS Clip tool to clip water bodies to the study OcnSt polygon
3.3.3. Used the ArcGIS Summary Statistics tool to sum the area of any water bodies
3.3.4. Calculate Land Area by subtracting water area from total area
3.3.5. Summed the water area values up to the study area level using the OcnSt-Study area cross reference published at http://transition.fcc.gov/wcb/iatd/neca.html.
3.3.6. Compute Pctwater: 100* Water Area / Total Area.

3.4. Calculating number of census blocks and total housing units

Total housing units is the sum of the housing units in a study area. Number of census blocks is total number of census blocks within the study area. These data come from 2010 census block data. A census block is associated with a study area if the census block's centroid is inside the study area. The census data and documentation are available at http://www.census.gov/geo/www/tiger/tgrshp2010/tgrshp2010.html. The calculation process used by Commission staff follows.

3.4.1. Used ArcGIS Feature to Point tool to create a point data set generated from the representative locations of 2010 census blocks. In that tool, Commission staff used the "inside" option to force the census block centroids to be inside the census block.
3.4.2. Used the ArcGIS Clip tool to clip census block centroids to the OcnSt polygon.
3.4.3. Retained the number of records that result from the clip tool as the number of blocks in the OcnSt polygon.
3.4.4. Used the ArcGIS Summary Statistics tool to sum Housing units (HOUSING10 attribute) of all Census Blocks to generate the number of housing units in the OcnSt polygon
3.4.5. Summed the values up to the study area level using the OcnSt-Study area cross reference published at http://transition.fcc.gov/wcb/iatd/neca.html.

3.5. Calculating PctUrban

Pcturban is the percentage of housing units in the study area that are in urban areas (the US Census Bureau's urbanized and urban clusters – URL below). "Urban Areas" layer comes from http://www2.census.gov/cgi-bin/shapefiles2009/national-files. Documentation on the urban areas can be found at http://www.census.gov/geo/www/ua/urbanruralclass.html. The calculation process used by Commission staff follows.

3.5.1. Used the ArcGIS Clip tool to clip urban areas to the OcnSt polygon
3.5.2. Used ArcGIS Feature to Point tool to create a point data set generated from the representative locations of 2010 census blocks. In that tool, Commission staff used the "inside" option to force the census block centroids to be inside the census block.
3.5.3. Used the ArcGIS Clip tool to clip census block centroids to urban area polygons (from section 3.5.1).

3.5.4. Used the ArcGIS Summary Statistics tool to sum Housing units (HOUSING10 attribute) of all Census Blocks to generate the number of housing units in urban area polygons.
3.5.5. Summed the values from all urban area polygons up to the study area level using the OcnSt-Study area cross reference published at http://transition.fcc.gov/wcb/iatd/neca.html.
3.5.6. Divided the sum by the number of the Total Housing units (section 3.4) in the study area.

3.6. Calculating Roadmiles

Roadmiles is the number of road miles within the study area. The calculation process is outlined below. All road types were included from the following data sets. With the exception of study areas in US territories, the street layer of the ESRI street maps were used for all study areas. Documentation: http://resources.arcgis.com/content/community-maps/street-map. For study areas in American Samoa and Guam, Commission staff used the roads layer from Census tiger files. The Tiger files are available at http://www.census.gov/cgibin/geo/shapefiles2010/main. The calculation process used by Commission staff follows.

3.6.1. Commission staff wrote a Python script to execute the following geoprocessing steps.
3.6.2. Used the ArcGIS Clip tool to clip road segments to the OcnSt polygon.
3.6.3. Used the ArcGIS function Calculate Geometry to calculate road length in miles.
3.6.4. Used the ArcGIS Summary Statistics tool to sum road miles of all road segments in the OcnSt polygon.
3.6.5. Summed the road miles up to the study area level using the OcnSt-Study area cross reference published at http://transition.fcc.gov/wcb/iatd/neca.html.

3.7. Calculating Roadcrossings

Roadcrossings is the number of road crossings within the study area. Each road intersection should have 3 or more road crossings. The calculation process used by Commission staff follows.

3.7.1. Commission staff wrote a Python script to execute the following geoprocessing steps.
3.7.2. Used the ArcGIS Feature Vertices to Points tool to calculate all potential road crossings in each OcnSt polygon.
3.7.3. Used the ArcGIS Intersect tool to calculate all potential road intersections.
3.7.4. Used the ArcGIS Frequency tool to calculate number of road crossings for each road intersection.
3.7.5. Saved road crossings of all real road intersections which have 3 or more road crossings.
3.7.6. Used the ArcGIS Clip tool to clip the road crossings to the OcnSt polygon.
3.7.7. Used the ArcGIS Summary Statistics tool to sum road crossings of all road intersections.
3.7.8. Summed the road crossings up to the study area level using the OcnSt-Study area cross reference published at http://transition.fcc.gov/wcb/iatd/neca.html.

3.8. Calculating Streamcrossings

Streamcrossings is the number of road-stream crossings within the study area. The NHDFlowline layer of the National Hydrography Dataset (NHD) was used for all study areas. Documentation: http://nhd.usgs.gov/index.html. The calculation process used by Commission staff follows.

3.8.1. Commission staff wrote a Python script to execute the following geoprocessing steps.

3.8.2. Used the ArcGIS Intersect tool to locate all road-stream crossings in each OcnSt polygon.
3.8.3. Used the ArcGIS Clip tool to clip the road-stream crossings to the OcnSt polygon
3.8.4. Used the ArcGIS Summary Statistics tool to compute total number of road-stream crossings.
3.8.5. Summed the stream crossings up to the study area level using the OcnSt-Study area cross reference published at http://transition.fcc.gov/wcb/iatd/neca.html.

3.9. Calculating Climate

Climate is the weighted average climate index (based on USDA's plant hardiness index) along the roads in the study area, weighted by the length of the road segment. For each road segment, determine the plant hardiness value, then use a look up table to identify the climate index value for the given hardiness value, and multiply climate index value by the length of the road. Then sum all these values and divide that sum by total road miles in the study area. The calculation process used by Commission staff follows.

3.9.1. Commission staff wrote a Python script to execute the following geoprocessing steps.
3.9.2. Used the ArcGIS Clip tool to clip road segments to the OcnSt polygon.
3.9.3. Used the ArcGIS Identity tool to compute the geometric intersections of the selected road segments and the hardiness zone to the OcnSt polygon. All the attributes from the road segments, as well as the hardiness zone information, were transferred to the output intersections.
3.9.4. Identified the climate effect value for the road segment based on the plant hardiness attribute "zone" using the ClimateIndex.xls file that is available at http://transition.fcc.gov/wcb/iatd/neca.html.
3.9.5. Multiplied the index value by the length of the road in miles.
3.9.6. Summed these values up to the study area level using the OcnSt-Study area cross reference published at http://transition.fcc.gov/wcb/iatd/neca.html.
3.9.7. Divided the sum by the number of road miles in the study area.

3.10. Calculating daysabvpt5

Daysabvpt5 is the average number of days in a year that the study area will receive more than 0.5 inches of rainfall. Weather station location data are available here: http://www1.ncdc.noaa.gov/pub/data/normals/1981-2010/station-inventories/allstations.txt. Data for the average number of days with rainfall amounts greater than 0.5 inches for individual weather stations (rainfall data set) is available at (http://www1.ncdc.noaa.gov/pub/data/normals/1981-2010/products/precipitation/ann-prcpavgnds-ge050hi.txt). The calculation process used by Commission staff follows.

3.10.1. Geolocated weather stations using the lat/long from the Weather station location data.
3.10.2. Each OcnSt polygon could have multiple single part polygons.
3.10.3. For each single part polygon, Commission staff used the ArcGIS Feature Vertices to Points tool to calculate the centroid of the polygon.
3.10.4. Used the ArcGIS function Calculate Geometry to calculate area of the single-part polygon.
3.10.5. Used the ArcGIS Clip tool to clip any weather stations to each single part polygon.
3.10.6. For those single part polygons that contained one or more stations, calculated the mean of the average number of days with rainfall above 0.5 inches from those stations using the rainfall data set.
3.10.7. If no station resided in the single part polygon, then commission staff used the ArcGIS Near tool to calculate distance between each and the centroid of the single-part polygon and then selected the rainfall value from the nearest station

3.10.8. For the multipart polygons, Commission staff calculated a weighted average for the OcnSt polygon as a whole using the area of the single part polygons as the weight.

3.11. Calculating pctwatertable36

Pctwatertable36 is the percentage of road segments where the water table is less than 36 inches below the surface of the ground along the roads in the study area. Pctwatertable36 is calculated by summing the road miles where the water table is less than 36 inches deep and dividing that sum by total road miles in the study area and then multiplying by 100. The U.S. General Soil Map (STATSGO2) was used for all study areas. The documentation is available at http://soils.usda.gov/survey/geography/statsgo/. The calculation process used by Commission staff follows.

3.11.1. Commission staff wrote a Python script to execute the following geoprocessing steps.

3.11.2. Used the ArcGIS Clip tool to clip road segments to the OcnSt polygon.

3.11.3. Used the ArcGIS Identity tool to compute geometric intersections of the road segments with soil polygons within the OcnSt polygon. All the attributes from the road segments, as well as the soil information, were transferred to the output intersections.

3.11.4. Obtained soil type of each road segment and retrieved its soil attributes.

3.11.5. Kept road segments where the "wtdepannmin" < 36 inches by extracting the value of the "wtdepannmin" field from the "muaggatt" table of the STATSGO2 database.

3.11.6. Summed the number of road miles calculated in the previous step up to the study area level using the OcnSt-Study area cross reference published at http://transition.fcc.gov/wcb/iatd/neca.html, and divide by the total number of road miles in the study area.

3.12. Calculating PctBedrock36

PctBedrock36 is the percentage of road segments with a study area that have bedrock within 36 inches of the surface of the ground along the roads. PctBedrock36 is calculated by summing the road miles where bedrock is within 36 inches of the surface and dividing that sum by the total road miles in the study area and then multiplying by 100. The U.S. General Soil Map (STATSGO2) was used for all study areas. Documentation: http://soils.usda.gov/survey/geography/statsgo/. The calculation process used by Commission staff follows.

3.12.1. Commission staff wrote a Python script to execute the following geoprocessing steps.

3.12.2. Used the ArcGIS Clip tool to clip road segments to the OcnSt polygon

3.12.3. Used the ArcGIS Identity tool to compute geometric intersections of the road segments with soil polygons within the OcnSt polygon. All the attributes from the road segments, as well as the soil information, were transferred to the output intersections.

3.12.4. Get soil type of a road segment and retrieve its soil attributes

3.12.5. Kept road segments where the "brockdepmin" < 36 inches by extracting the value of the "brockdepmin" field from the "muaggatt" table of the STATSGO2 database

3.12.6. Summed the number of road miles calculated in the previous step up to the study area level using the OcnSt-Study area cross reference published at http://transition.fcc.gov/wcb/iatd/neca.html, and divided by the total number of road miles in the study area.

3.13. Calculating Slope

Slope is the weighted average slope along the roads in the study area, weighted by the length of the road segment. For each road segment, Commission staff used the slope of that road segment and multiplied it by the length of the road in miles. Then staff summed those values and divided that sum by total road miles in the study area. The U.S. General Soil Map (STATSGO2) was used for all study areas. Documentation: http://soils.usda.gov/survey/geography/statsgo/. The calculation process used by Commission staff follows.

3.13.1. Commission staff wrote a Python script to execute the following geoprocessing steps
3.13.2. Used the ArcGIS Clip tool to clip road segments to the OcnSt polygon
3.13.3. Used the ArcGIS Identity tool to compute geometric intersections of the road segments with soil polygons within the OcnSt polygon. All the attributes from the road segments, as well as the soil information, were transferred to the output intersections.
3.13.4. Obtained soil type of each road segment and retrieved its soil attributes
3.13.5. Extracted the "slopegradwta" field from the "muaggatt" table of the STATSGO2 database
3.13.6. Multiplied slopegradwta by the length of the road segment.
3.13.7. Summed the numbers calculated in the previous step up to the study area level using the OcnSt-Study area cross reference published at http://transition.fcc.gov/wcb/iatd/neca.html, and divided by the total number of road miles in the study area.

3.14. Calculating (soil construction) Difficulty

Difficulty is the weighted average soil construction difficulty along the roads in the study area, weighted by the length of the road segment. For each road segment, calculate the soil construction difficulty. Calculate difficulty value based on the dominant soil texture using the SoilTextureStatsgo2Index.xls file, which is available at http://transition.fcc.gov/wcb/iatd/neca.html and multiply that value by the length of the road segment to get an intermediate value. Then sum all these intermediate values and divide that sum by total road miles in the study area. The U.S. General Soil Map (STATSGO2) was used for all study areas. Documentation: http://soils.usda.gov/survey/geography/statsgo/. The calculation process used by Commission staff follows.

3.14.1. Commission staff wrote a Python script to execute the following geoprocessing steps
3.14.2. Used the ArcGIS Clip tool to clip road segments to each OcnSt polygon
3.14.3. Used the ArcGIS Identity tool to compute geometric intersections of the road segments with soil polygons within the OcnSt polygon. All the attributes from the road segments, as well as the soil information, were transferred to the output intersections.
3.14.4. Found the dominant component with the highest percentage of the soil type (from the "comppct" field from the "muaggatt" table of the STATSGO2 database);
3.14.5. Found the thickest horizon within that dominant component that is within the top 36 inches of the soil
3.14.6. Found the dominant soil texture for that (thickest) horizon.
3.14.7. Calculated difficulty value based on the dominant soil texture using the SoilTextureStatsgo2Index.xls file, which is available at http://transition.fcc.gov/wcb/iatd/neca.html.
3.14.8. Summed the numbers calculated in the previous step up to the study area level using the OcnSt-Study area cross reference published at

http://transition.fcc.gov/wcb/iatd/neca.html, and divided by the total number of road miles in the study area.

3.15. Calculating PctTribalLand

PctTribalLand is the study area's percentage of land area that consists of federal tribal lands. The data are available at http://www.census.gov/cgi-bin/geo/shapefiles2010/main. The calculation process used by Commission staff follows.

3.15.1. Used the Select by Location function to select census block centroids in the study area (calculated in step 3.4.1) that intersect census tribal areas

3.15.2. Programmatically linked those census blocks with census data containing the land area for each block.

3.15.3. For each OcnSt polygon, calculated the sum of the land area of the blocks inside the federal tribal lands.

3.15.4. Summed the values calculated in the previous step up to the study area level using the OcnSt-Study area cross reference published at http://transition.fcc.gov/wcb/iatd/neca.html, divided by the study area's total land area, and then multiplied by 100.

3.16. Calculating PctParkLand

PctParkLand is the study area's percentage of land area that consists of national park land. Data available at http://www.bts.gov/programs/geographic_information_services/. The calculation process used by Commission staff follows.

3.16.1. Used the Select by Location function to select census block centroids (calculated in step 3.4.1) that intersect national park land area.

3.16.2. Programmatically linked the census blocks with census data containing the land area for each block.

3.16.3. For each OcnSt polygon, calculated the sum of the land area of the blocks inside the park lands.

3.16.4. Summed the values calculated in the previous step up to the study area level using the OcnSt-Study area cross reference published at http://transition.fcc.gov/wcb/iatd/neca.html, divided by the study area's total land area, and then multiplied by 100.

4. Study Area Boundary Edits

The previously detailed algorithms should be sufficient to replicate the Commission's geospatial variable values used in the HCLS benchmark regressions for those study areas that do not have special circumstances described below. For those study areas described below, Commission staff examined public Commission sources and publicly available data to create or modify the TeleAtlas boundaries. Because this was a manual process, it could be more difficult for others to precisely replicate the variable values for these study areas. The process used by Commission staff follows.

4.1. Missing study area boundaries/OcnSt polygons

4.1.1. Data for some of the missing study areas with missing boundaries in the September 2011 data were obtained from earlier versions of the TeleAtlas data.

4.1.2. For those study areas for which we could not find any boundaries in any of the versions of the TeleAtlas data that the Commission obtained, other public sources of data were used to manually create them, such as the tribal land that the study area serves or study area waivers.

4.2. Study areas receiving frozen support

4.2.1. For each study area that receives frozen support, the area associated with the frozen support needed to be separated from the rest of the study area's OcnSt polygon so that the geospatial variables could be based on the non-frozen support area. Commission staff examined public Commission sources and publicly available data and manually excluded these areas from these OcnSt polygons.

5. Variables created using NECA Overview data
   5.1. LnLoops is the natural log of the 2010 DL060 loop count.
   5.2. PctLoopChange is the percentage change of DL060 loop count between 2009 and 2010. For the observations that converted from being average schedule to cost companies (and therefore did not have DL060 loop counts for the prior year), used the percentage change in DL070 loops.
   5.3. PctUndepPlant is the percentage of the plant that has not yet been depreciated. It is 100 * DL220 / DL160 (i.e., 100*net plant/gross plant).
   5.4. LnExchanges is the natural log of the number of exchanges in the study area.

6. Variables using other information
   6.1. StateSACs is the number of study areas in the state that are owned by the same holding company or have common control in the state. LnStateSacs is the natural log of StateSacs. The holding company/common control ownership information can be found in the Universal Service Monitoring Report, CC Docket No. 98-202, app. (2011) (HC NECA ILEC Support Data - by Study Area.xls), available at http://www.fcc.gov/Bureaus/Common_Carrier/Reports/FCCState_Link/Monitor/2011_MR_Supplementary_Material.zip.
   6.2. Census regions were determined by the primary state of the study area. If the study area's state is in the region, then the study area's region variable is a 1, else a 0. Census regions were determined by the primary state of the study area. If the study area's state is in the region, then the study area's region variable is a 1, else a 0. The list of states in each region follows.
      6.2.1. Western (AK, CA, CO, HI, ID, MT, NM, NV, OR, UT, WA, WY)
      6.2.2. Midwest (IA, IL, IN, KS, MI, MN, MO, ND, NE, OH, SD, WI)
      6.2.3. Northeast (CT, MA, ME, NH, NJ, NY, PA, RI, VT,)
      6.2.4. South (AL, AR, AZ, DC, DE, FL, GA, KY, LA, MD, MS, NC, OK, SC, TN, TX, VA, WV)
   6.3. Alaska is a 1 if the study area is located in Alaska, else a 0.
   6.4. Medhomeval2000 is the Median Home value in the state in 2000. It is available at http://www.census.gov/hhes/www/housing/census/historic/values.html. LnMedhomeval2000 is the natural log of Medhomeval2000.
   6.5. Density is the number of housing units in a study area (see step 3.4) divided by the size of the study area in square miles (see step 3.2). LnDensity is the natural log of density.
   6.6. Wdensity is the weighted density of the study area. For each census block in a study area, calculate density by dividing housing units by land area. Then calculate the weighted average for the study area using the housing units as the weight. Lnwdensity is the natural log of weighted density.

Appendix B: Analysis of Dependent Variables (Capex and Opex)

The designers of the QRA appear to have incorrectly defined the costs associated with the dependent variables. The goal of the QRA was to establish benchmarks for costs to deploy and operate the high-cost loop networks. As explained earlier, the QRA should identify the answers to two questions.

- How much subscriber cable and wire facility loop plant and subscriber circuit equipment is needed to provide universal services in comparison to the costs of other companies?
- How much does it cost to operate the subscriber portion of the network in comparison to the costs at other companies?

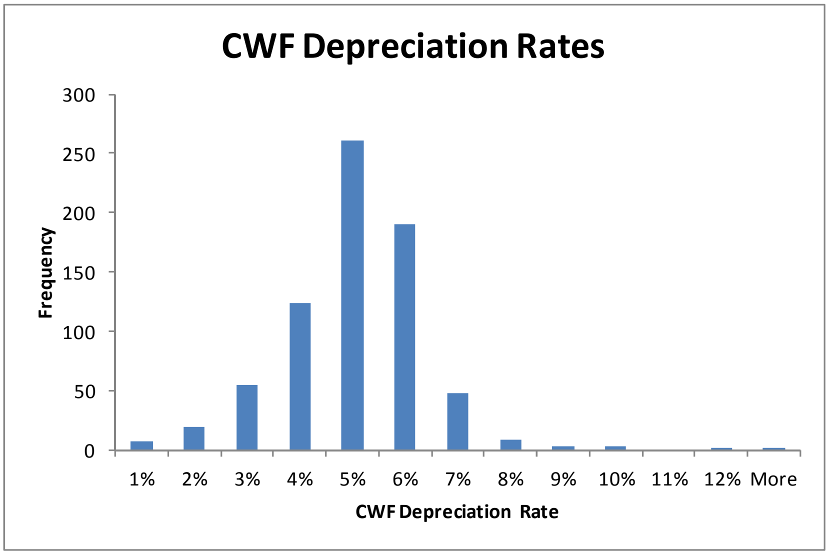
The FCC uses costs developed by the NECA HCL Algorithm to calculate Study Area Cost per Loop, a process that uses allocated or derivative costs rather than actual costs. The allocation percentages vary greatly from company to company, making the comparisons problematic. Further, the input data are out of the companies' control since they are established by Code of Federal Regulation rules or state public utility commission orders. The allocations introduce an unpredictable, complicating factor that degrades or invalidates the relationships between the predictor variables and the dependent variables (opex and capex).

CAPEX SHOULD NOT BE BASED ON DEPRECIATION EXPENSE

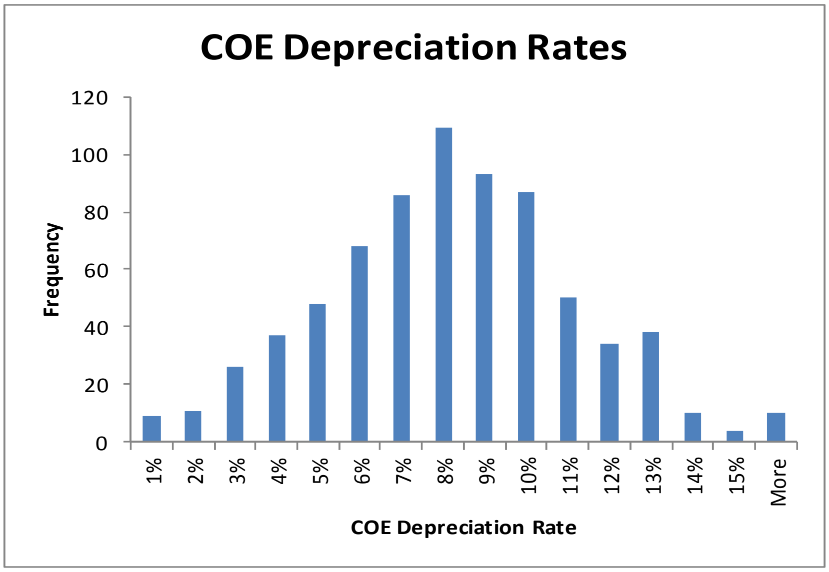
Instead of using the actual asset costs of cable and circuit equipment, the FCC defined the capital costs (capex) as depreciation expense of subscriber cable and wire facility ("CWF") and subscriber circuit equipment plus the return on investment component.

While there is a cost-causative relationship between certain variables, such as the number of road crossings, in a study area and cost of CWF deployment in the local network, there is not a cost-causation between the number of road crossings and the depreciation expense based on a rate determined differently by the various state public utility commissions. In fact, there is not and cannot be a cost relationship between any variable and an arbitrary number.

The histogram below shows the widely-divergent depreciation rates for cable and wire facilities drawn from the companies' cost data used in the QRA[54].



In fact, cable and wire facility depreciation rates range from less than 2% to over 10% annually. Even if one only reviewed the highest frequency of rates between 4% and 7%, the variance is more than 50% from 4% to 7%. The situation is even more extreme for subscriber equipment, where depreciation rates range from less than 2% to more than 15% annually.



54 Because mass asset depreciation is used by the majority of local exchange carriers, we can calculate the effective depreciation rates of asset classes from the HCL Data Input provided by the FCC. Cable and Wire Facility depreciation was calculated as Data Line 530: Depreciation Expense- CWF divided by Data Line 700: Cost Study Average CWF from the FCC QRA Input File *RegInput2012_0425.xls* provided at http://transition.fcc.gov/wcb/iatd/neca.html.

It is helpful to examine an illustration to see how the improper inclusion of an arbitrary number without cost causation (in this case the depreciation rate) undermines the validity of the relationships. Suppose we have two identical companies. Both Company A and Company B have the same number and density of customers; same size, shape and make-up of service area; the same amount of local loop plant deployed, etc. Everything is exactly the same except their depreciation rates where Company A's rate is 4% and Company B's rate is 6%. Based on this identical make-up, one should be able to take a variable value from one company and predict the cost of the other company if cost causation exists.

Returning to the cost relationship between the number of road crossings and the cost of CWF, it is obvious that there is an identical cost relationship.

	Company A		Company B
Cable & Wire Facility Assets	$ 10,000,000		$ 10,000,000
Road Crossings	250		250
CWF $ per Road Crossing	$ 40,000		$ 40,000

However, when arbitrary, non-cost-causative factors such as depreciation rates are introduced, the relationship between predictor and cost is broken in spite of the fact that the companies are otherwise identical.

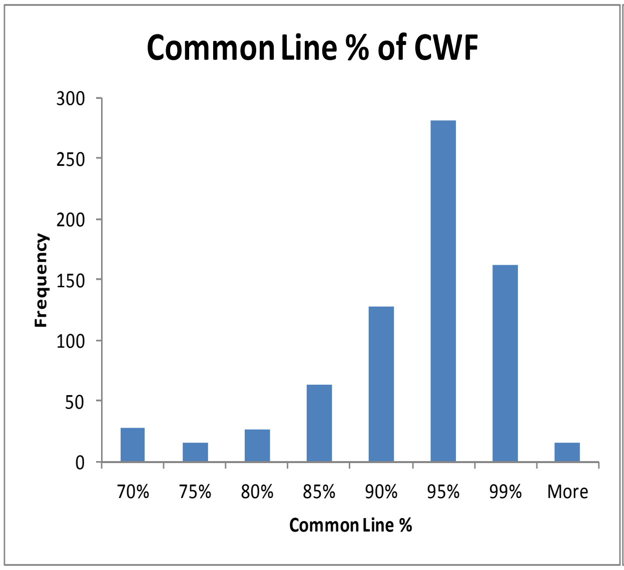
	Company A		Company B
Cable & Wire Facility Assets	$ 10,000,000		$ 10,000,000
Depreciation Rate	4%		6%
Depreciation Expense	$ 400,000		$ 600,000
Road Crossings	250		250
Depreciation Expense per Road Crossing	$ 1,600		$ 2,400

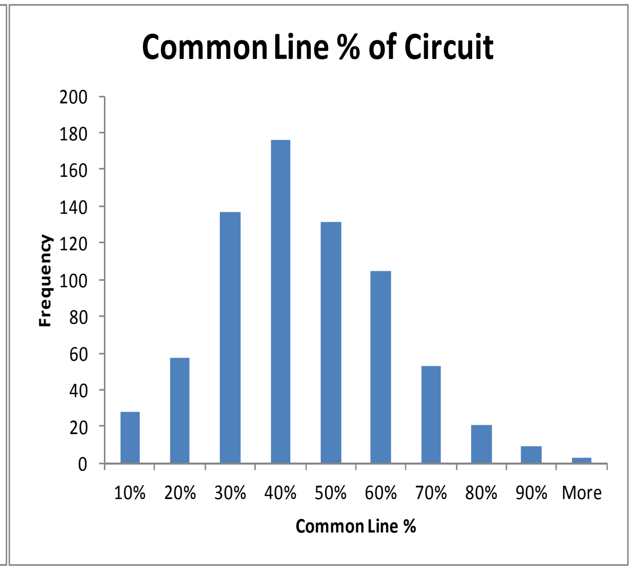
The use of depreciation expense as a primary component of capex is improper because there is no cost-causation between the depreciation rate and the cost predictor variables. It distorts or invalidates the results of the Quantile Regression Analysis.

OPEX SHOULD NOT BE BASED ON ALGORITHM COSTS ALLOCATED TO COMMON LINE

The Bureau similarly defined Operating Costs (opex) as the sum of NECA High Cost Loop Algorithm steps 13, 14, 15, 16, 20, 21, and 22. The costs in the algorithm are not the actual costs from the company's income statement, but rather expenses allocated to common line usage. Consequently, the development of cost relationships for opex is distorted by the differences in company common line allocations, which occur differently for the various companies evaluated by the QRA. These common line allocation amounts are determined by network architecture needs and 47 CFR Part 36 Separations Rules and are largely out of each company's control.

The allocation differences might be overlooked if they were immaterial but as the following histograms show, there is substantial variation in the allocations. The common line percentage for Circuit is particularly broad.



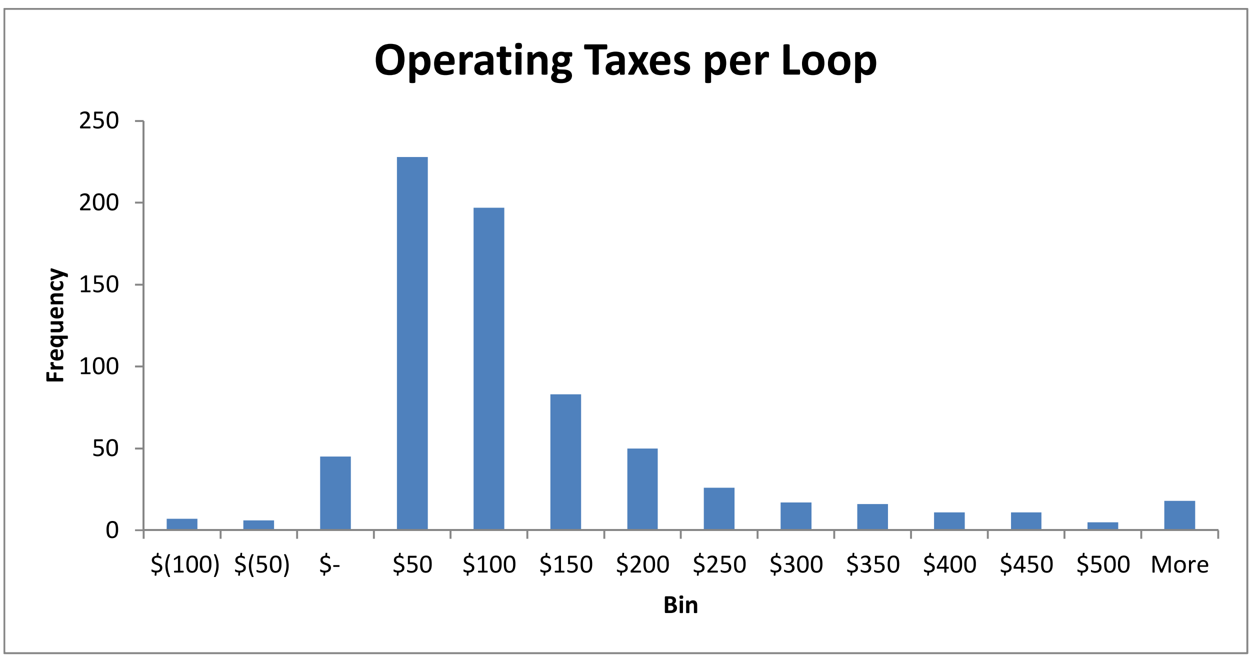


Defining opex as the sum of allocated expenses is improper because there is no cost causation between the allocation percentages and the cost predictor variables. A company may appear to have lower costs when in reality it is common line allocation percentages that falsely skew the results. The wide variance in allocations distorts the Quantile Regression Analysis.

Appendix C: Analysis of Operating Taxes in Opex

Operating taxes include property taxes, operating investment tax credits, state and federal income taxes, deferred operating taxes, and other operating taxes. Like the inclusion of depreciation rates in capex, the inclusion of arbitrary tax rates undermines the establishment of cost relationships. No variable used by the FCC for scale, geography, customer dispersion, etc. can predict that the state of Nevada has 0% corporate income tax rates while Pennsylvania has a 9.99% corporate income tax rate.[55] Furthermore, income tax rates are dependent on corporate structure and many cooperatives pay no income taxes while a privately-owned company may pay as much as 39%. Once again, these costs are wholly out of the control of the companies. If a company is paying "too much" in operating taxes in comparison to other companies, there is no way for the company to control that cost. The inclusion of Operating Taxes in the benchmarking process is completely contrary to the goal of any regression analysis which is to identify valid relationships. The inclusion of Operating Taxes dramatically skews the results of the QRA and unfairly punishes the customers of companies that are privately-owned or are located in states with higher property or income tax rates.

The authors compared the amount of operating taxes for the 726 company study areas in the QRA in two different manners.[56] First, we compared operating tax expense per loop since the number of loops has the highest identified correlation with opex. Operating tax expense per loop ranged from ($1,035) per loop to $3,675 per loop. In fact, 58 of the 726 study areas had zero or negative amounts recorded for operating taxes expense while 61 had operating taxes per loop in excess of $300 per loop. The following histogram shows the frequency distribution.



[55] State corporate income tax rates were obtained from the Federation of Tax Administrators at http://www.taxadmin.org/fta/rate/corp_inc.pdf

[56] Data Line 650: Operating Taxes in the FCC QRA Input File *RegInput2012_0425.xls* provided at http://transition.fcc.gov/wcb/iatd/neca.html.

We also compared operating taxes to the total telephone plant in service ("TPIS") amounts of the carriers.[57] In this case, operating taxes as a percentage of TPIS ranged from -3.42% to 9.33%. Once again no discernible pattern could be detected to account for the wide disparity of values in the data.

Operating tax expense is a function of taxable income, expenses, corporate structure, federal tax rates, state tax rates, depreciation rates, property value (for property taxes), Internal Revenue Service Rules, and many other factors. The critical insight is that no independent variable (alone or in combination) used or contemplated for the QRA can predict operating taxes expense. Therefore, it is improper to include operating taxes in the regression analysis.

Furthermore, the inclusion of operating taxes in the benchmarking process frustrates the FCC's goal of rewarding operating "efficiency" as it distorts the QRA results. The inclusion of operating taxes in the definition of opex improperly punishes companies required to pay higher taxes while improperly rewarding companies that may not be operating as efficiently but have a lower tax burden. It also seems perverse and contrary to good public policy to restrict universal service support funds to companies because they pay comparatively more state and federally required taxes.

A re-creation of the QRA demonstrates the materiality of this issue. In the FCC's April 25, 2012 QRA, 63 study areas were over the opex cap. Alexicon ran an alternate QRA which excluded operating taxes from the definition of opex. The alternate analysis recalculated the 90th quantile benchmarks for opex less operating taxes based on the sixteen independent variables, and calculated revised caps for all 726 study areas in the QRA. Under this scenario, 17 of the original 63 opex-capped companies were under the opex less operating taxes cap while 18 previously uncapped companies were now above the benchmark. The number of capped companies alone does not capture the magnitude of the impact of benchmarked operating taxes on support. Of the 726 study areas, 585 qualify for high cost loop support according to the FCC's original calculations. The change in benchmark methodology described above would have resulted in changed high cost loop support amounts for 80 of the 585 companies ranging from losses of $725,000 to gains of $7.6 million annually. We respectfully submit that a process that uses non-comparable data to take $7.6 million in annual support from a company based on that company's obligation to pay relatively higher taxes is a flawed process in need of revision.

With due respect, the FCC's justification that operating taxes should be included in the benchmarking process because "taxes are an expense that must be paid, just like other operational expenses"[58] is insufficient reason to merit inclusion in the benchmarking process. The facts remain that operating taxes cannot be predicted by the QRA and the costs themselves are out of the control of the companies. The inclusion of operating taxes in the opex amount arbitrarily rewards and punishes companies and is contrary to the FCC's stated goals. Operating tax amounts should not be included in the calculation of operating expense limits but rather should be included in the calculation of recoverable costs *after* any limits are imposed.

[57] Data Line 210: Total Telecommunications Plant in Service in the FCC QRA Input File *RegInput2012_0425.xls* provided at http://transition.fcc.gov/wcb/iatd/neca.html.
[58] HCLS Benchmark Order, Appendix A, paragraph 79.

Appendix D: Scatterplots of FCC Independent Variables

The scatterplots in this Appendix are graphical depictions of the relationship between, on the one hand, capex, opex and the alternate cost recommended by the authors and, on the other hand, each of the independent variables used in the QRA. Visual inspection of the data is a helpful step in testing the statistical outputs. Data with higher correlation will appear as a tighter grouping heading in a discernible direction. Data with low correlation appear more dispersed.

> *"One should never base important conclusions on the value of the correlation coefficient alone (i.e., examining the respective scatterplot is always recommended)."*[59]

Each of the following pages presents three scatterplots that graph the natural log for each of the independent variables (except in instances where the independent variable is a percentage or "dummy" variable) and the natural log of the defined cost as follows:

BLUE: FCC Capex [60]
RED: FCC Opex[61]
GREEN: Alternate Cost [62]

The authors' recommended Alternate Cost combines capex and opex into a single cost amount and removes arbitrary and allocated amounts as discussed in Appendices B & C, that is, combining the two dependent variables, excluding operating taxes from the regression, and redefining capital expenditures as the reported asset amount. As is apparent in the graphs, the alternate definition of capital and operating

---

[59] StatSoft Electronic Statistics Textbook, https://www.statsoft.com/textbook/basic-statistics/#Correlationsa.

[60] FCC Capex = Depreciation & Amortization Expense - CWF Cat 1 (AL17)
+ Depreciation & Amortization Expense - COE Cat 4.13 (AL18)
+ Return Component - CWF Cat 1 (AL23)
+ Return Component - COE Cat 4.13 (AL24)

[61] FCC Opex = Maintenance Expense – Allocated to CWF Cat 1 (AL13)
+ Maintenance Expense – Allocated to COE Cat 4.13 (AL14)
+ Network & General Support Expense – Allocated to CWF Cat 1 & COE Cat 4.13 (AL15)
+ Network Operations Expense – Allocated to CWF Cat 1 & COE Cat 4.13 (AL16)
+ Corporate Operations Expense – Allocated to CWF Cat 1 & COE Cat 4.13 (AL19)
+ Operating Taxes – Allocated to CWF Cat 1 & COE Cat 4.13 (AL20)
+ Benefits – Allocated to CWF Cat 1 & COE Cat 4.13 (AL21)
+ Rents – Allocated to CWF Cat 1 & COE Cat 4.13 (AL22)
"ALxx" refers to the algorithm line of the NECA High Cost Loop Support algorithm

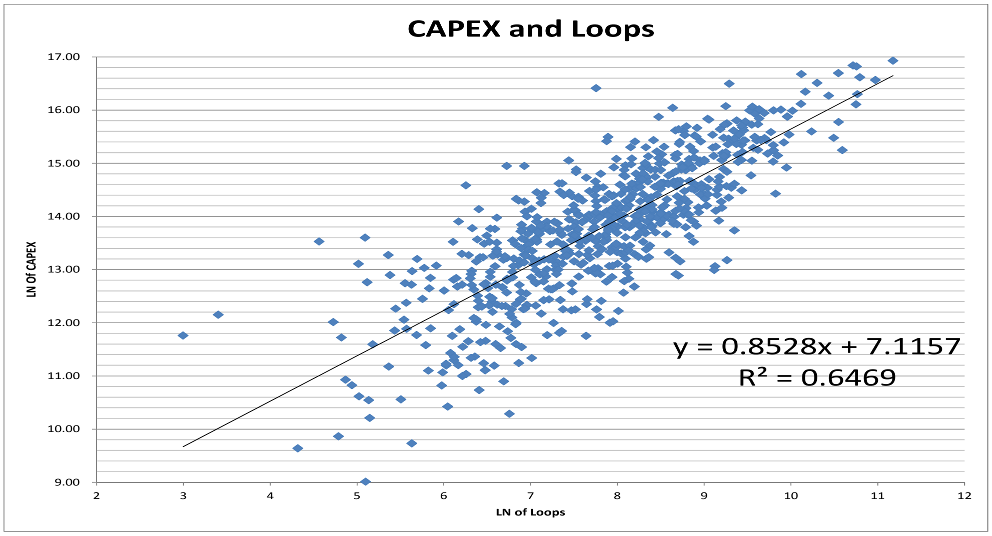
[62] Alternate Cost = Cable & Wire Facility Cat 1 Asset Balance (AL 1)
+ Central Office Equipment Cat 4.13 Asset Balance (AL 2)
+ Maintenance Expense – CWF (DL430 - DL435 - DL440)
+ Maintenance Expense – COE (DL365 + DL380 + DL395 - DL370 - DL375 - DL385 - DL390 - DL400 - DL405)
+ Network & General Support Expense (DL335 + DL350 - DL340 - DL345 - DL355 - DL360)
+ Network Operations Expense (DL450 - DL455)
+ Corporate Operations Expense (DL535 + DL550)
+ Benefits (DL600 - DL540 - DL555)
+ Rents (DL610)
"DLxx" refers to the data entry line of the NECA High Cost Data Input Form

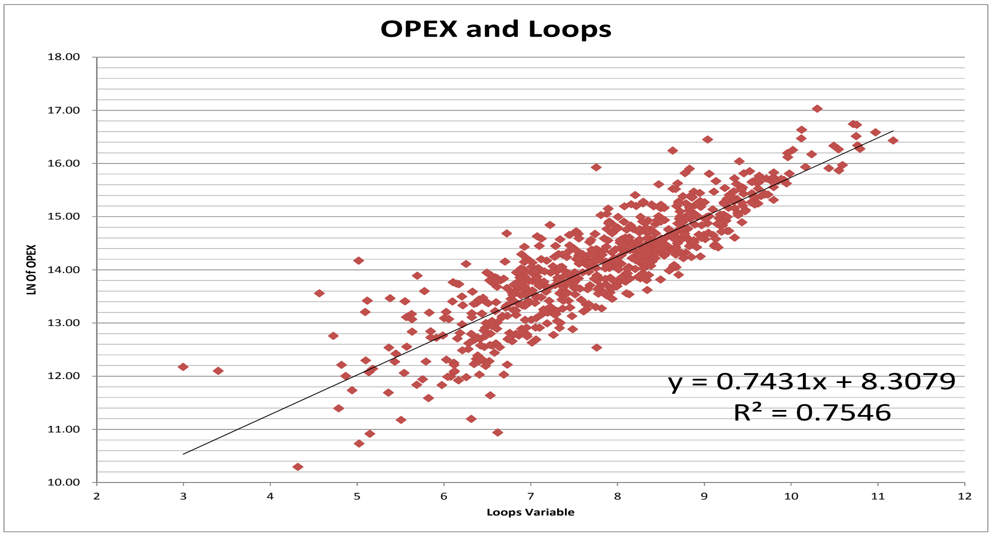
costs provides better correlation in most cases even before addressing the problems with the independent variables noted earlier.[63]

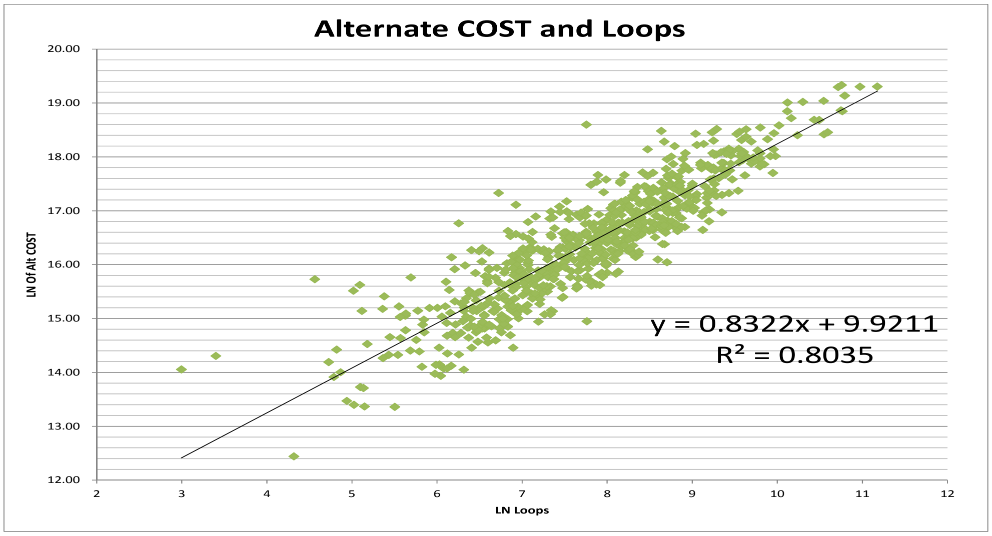
More significant is the fact that, out of the entire set of sixteen variables, the R^2 for the alternate regressions was superior to both the capex and opex regressions for all four of the variables—loops (0.8035), road crossings (0.5742), exchanges (0.4288), and road miles (0.4166)—for which the R^2 was higher than 0.1023. The alternative regression did not match the R^2 for capex and opex regressed on variables that returned very low R^2, which were Alaska (0.0007 for capex, 0.0012 for opex, 0.0002 for the alternate), percentage bedrock (0.0038 for capex, 0.0027 for opex, 0.0019 for the alternate), and number of study areas in the state (0.0209 for capex, 0.0154 for opex, 0.0119 for the alternate). We would argue that the alternate regression was superior for capex and opex based on each of the independent variables that proved more significant compared with the other variables.

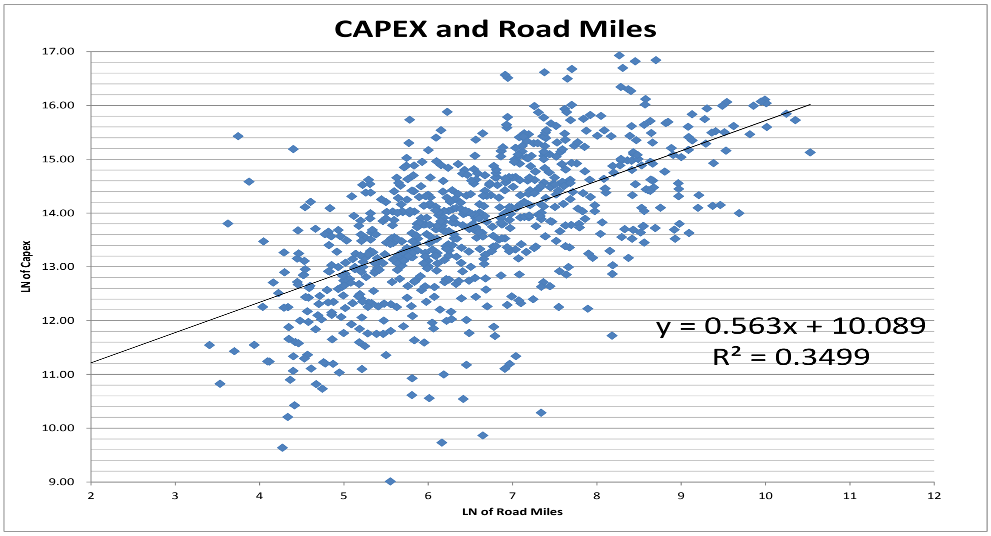
[63] The alternate regression R^2 was superior to that of the capex and opex calculations for 5 variables (loops, road miles, road crossings, and soil difficulty), was below the capex and opex for 3 variables (Alaska, number of study areas in the state, and bedrock, and was second best (below either capex or opex) for the remaining 8; so the alternative, without refining other variables, was below only 14 of the total 32 FCC regressions.

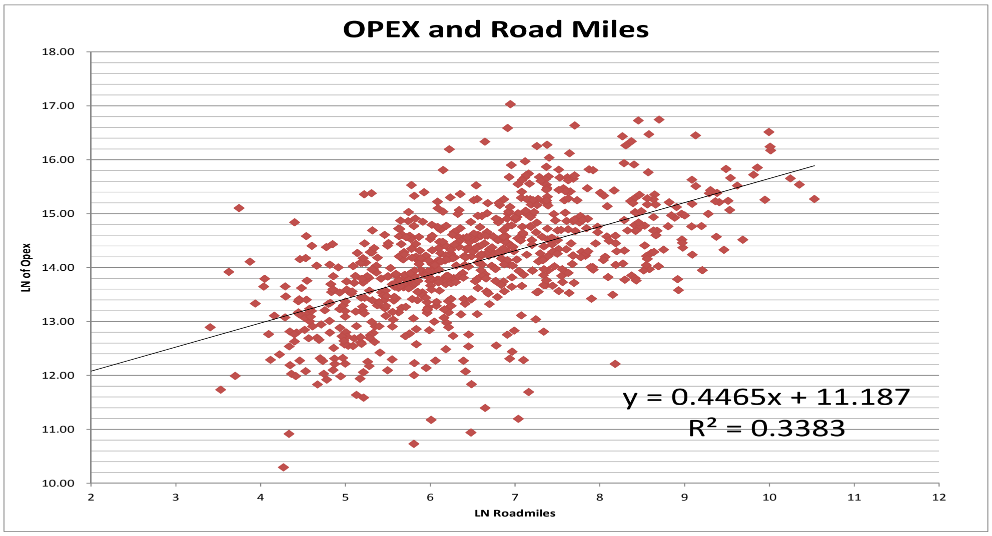
SCATTERPLOTS:

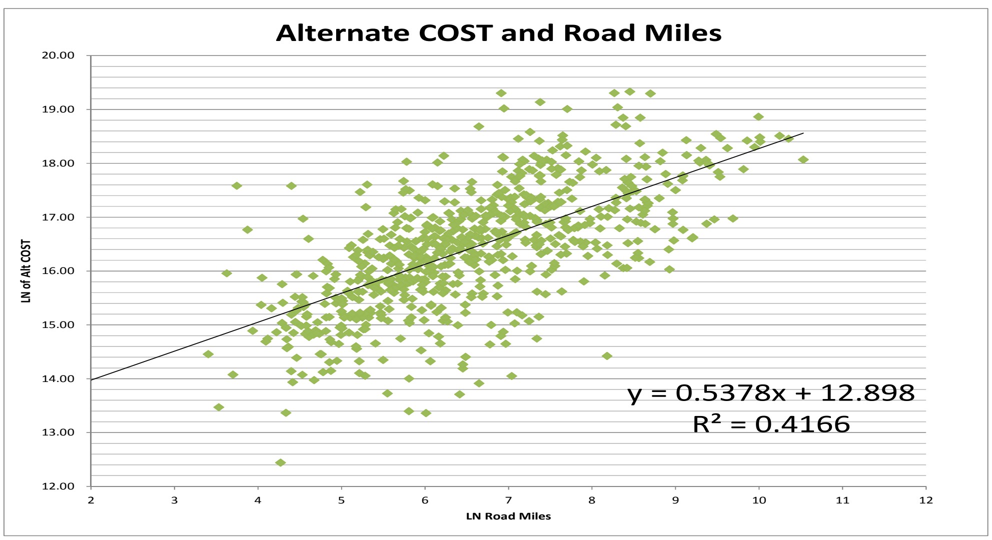


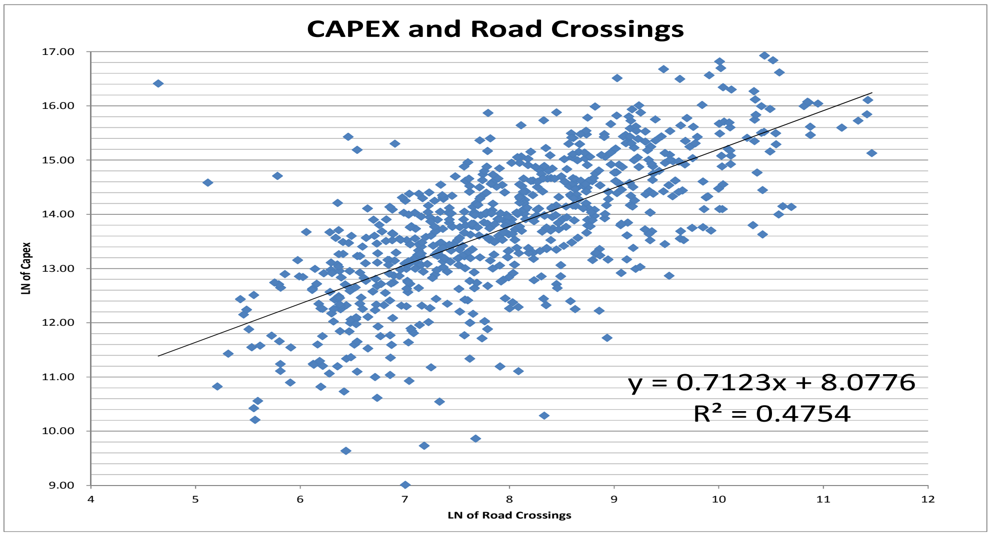


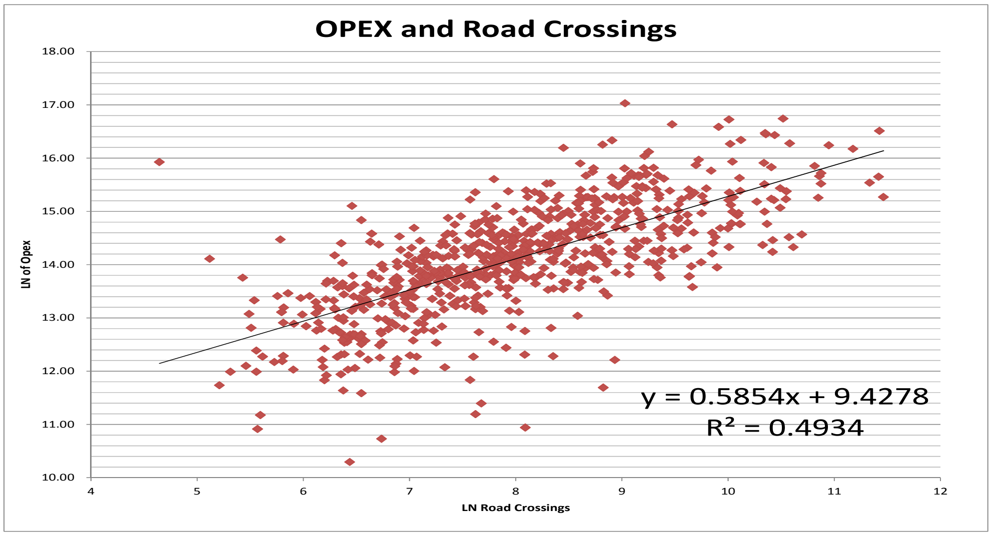


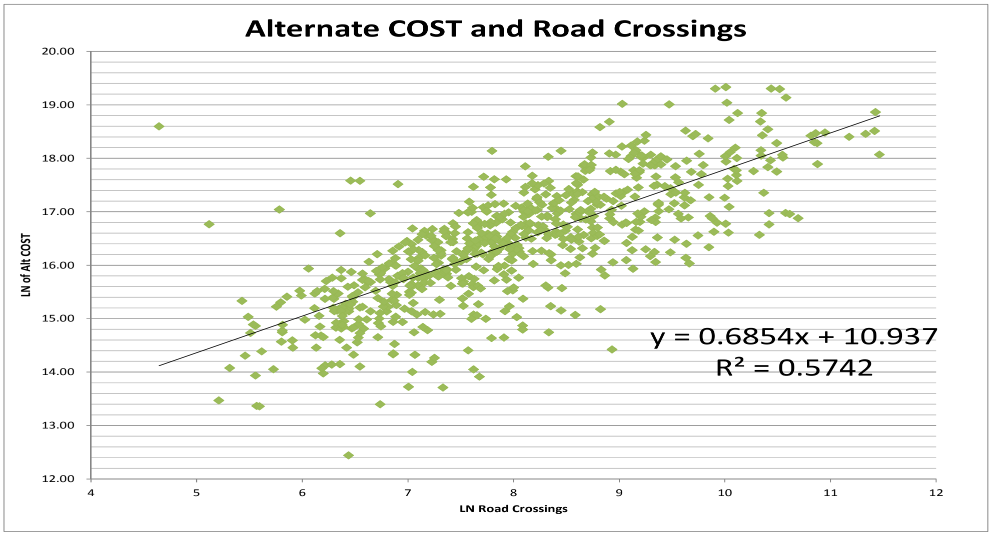

CAPEX and Road Miles
LN of Capex
LN of Road Miles
y = 0.563x + 10.089
R² = 0.3499

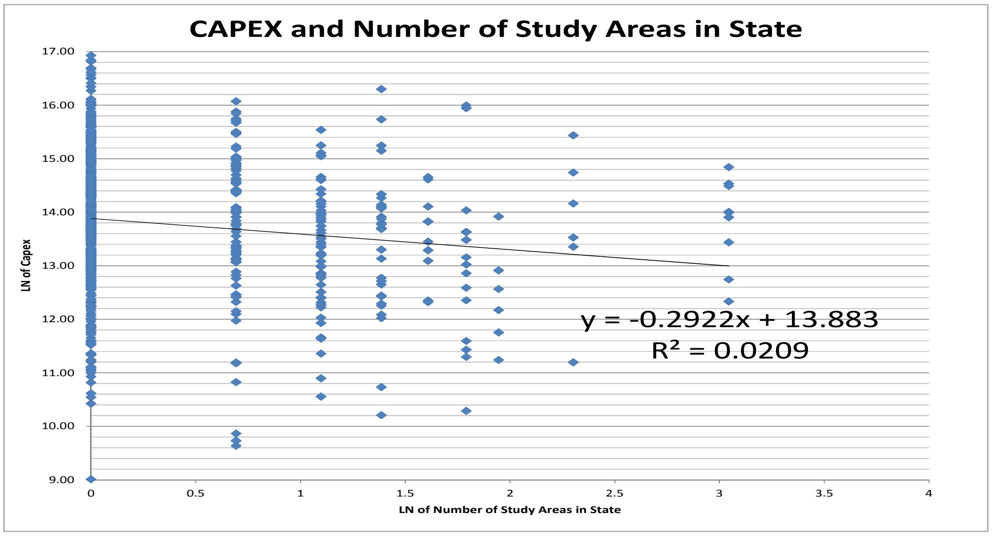

OPEX and Road Miles
LN of Opex
LN Roadmiles
y = 0.4465x + 11.187
R² = 0.3383

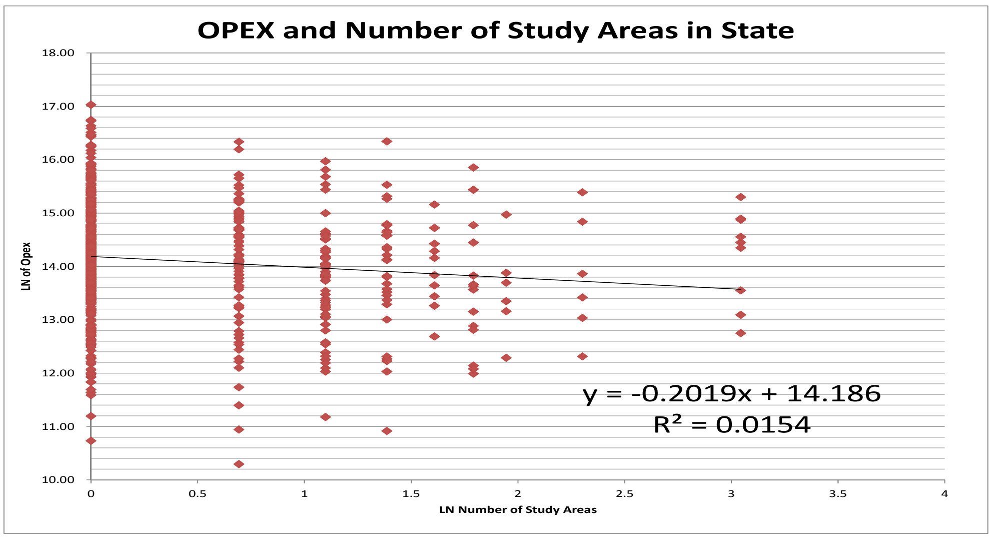

Alternate COST and Road Miles
LN of Alt COST
LN Road Miles
y = 0.5378x + 12.898
R² = 0.4166

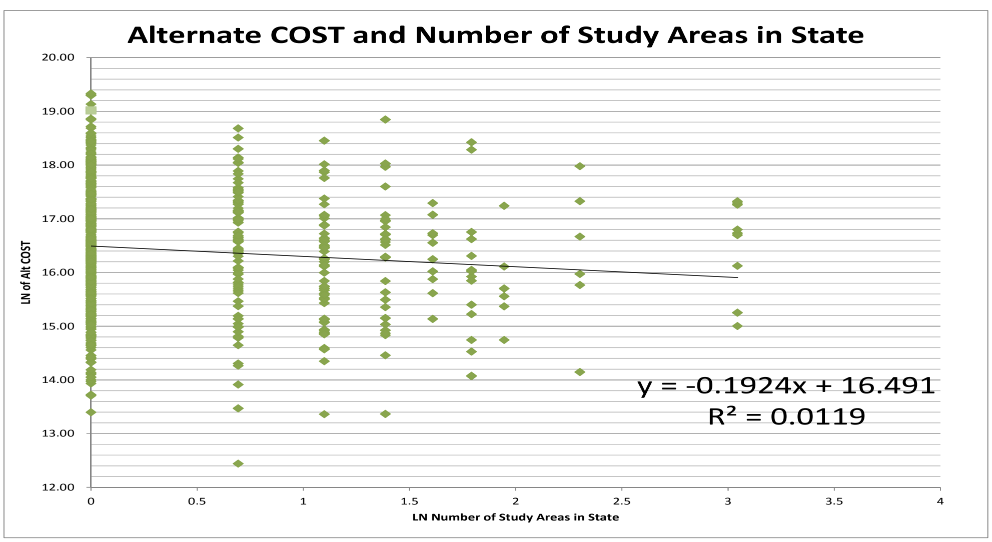

CAPEX and Road Crossings
LN of Capex
LN of Road Crossings
y = 0.7123x + 8.0776
R² = 0.4754

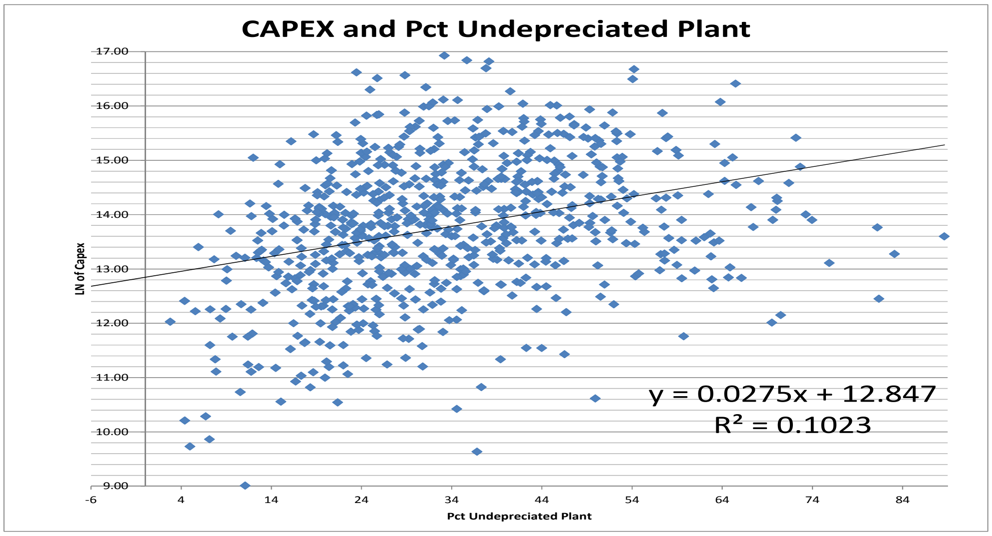

OPEX and Road Crossings
LN of Opex
LN Road Crossings
y = 0.5854x + 9.4278
R² = 0.4934

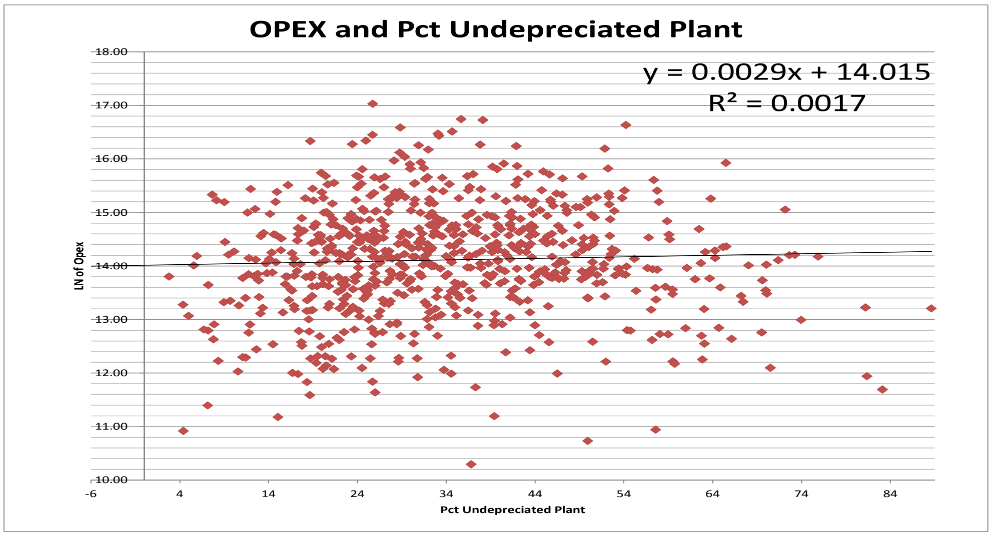

Alternate COST and Road Crossings
LN of Alt COST
LN Road Crossings
y = 0.6854x + 10.937
R² = 0.5742

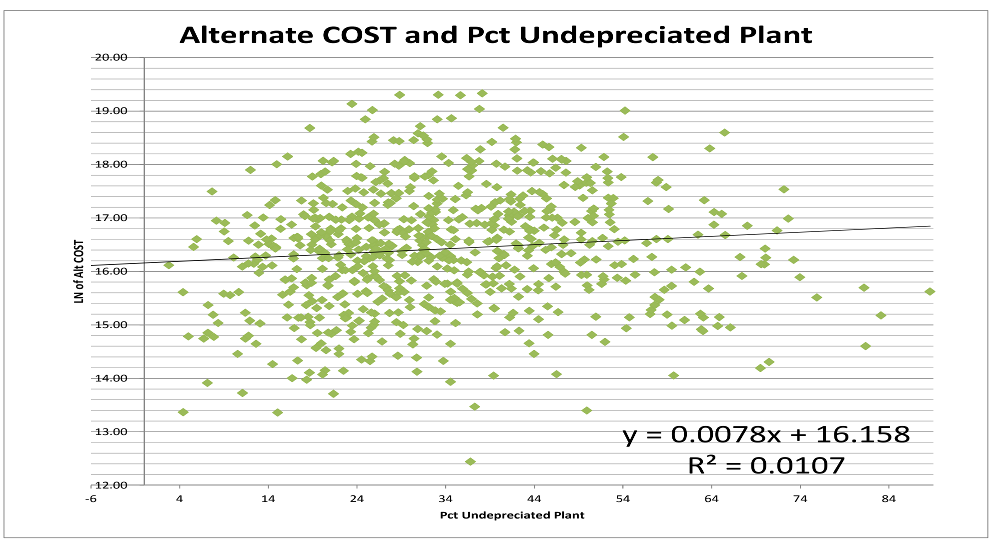

CAPEX and Number of Study Areas in State
LN of Capex
LN of Number of Study Areas in State
y = -0.2922x + 13.883
R² = 0.0209

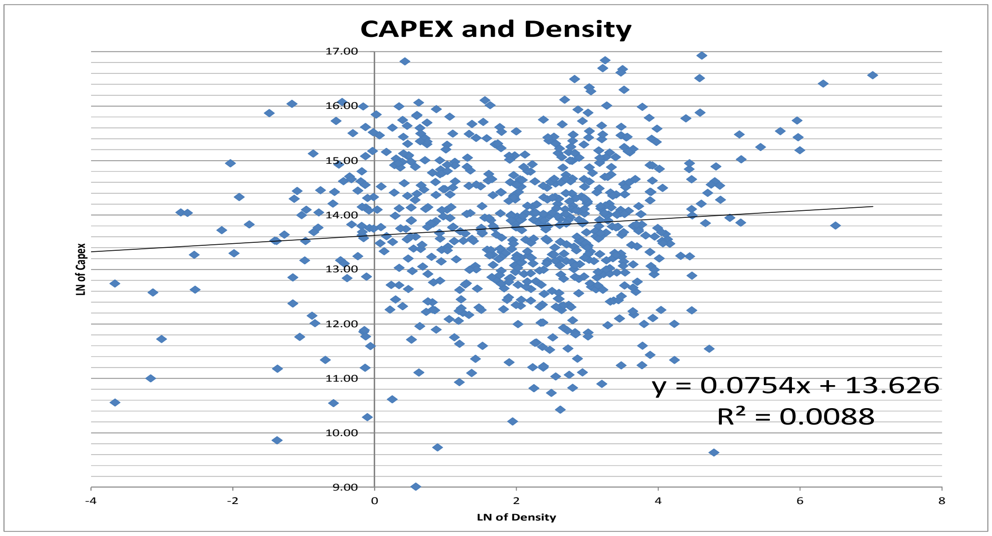

OPEX and Number of Study Areas in State
LN of Opex
LN Number of Study Areas
y = -0.2019x + 14.186
R² = 0.0154

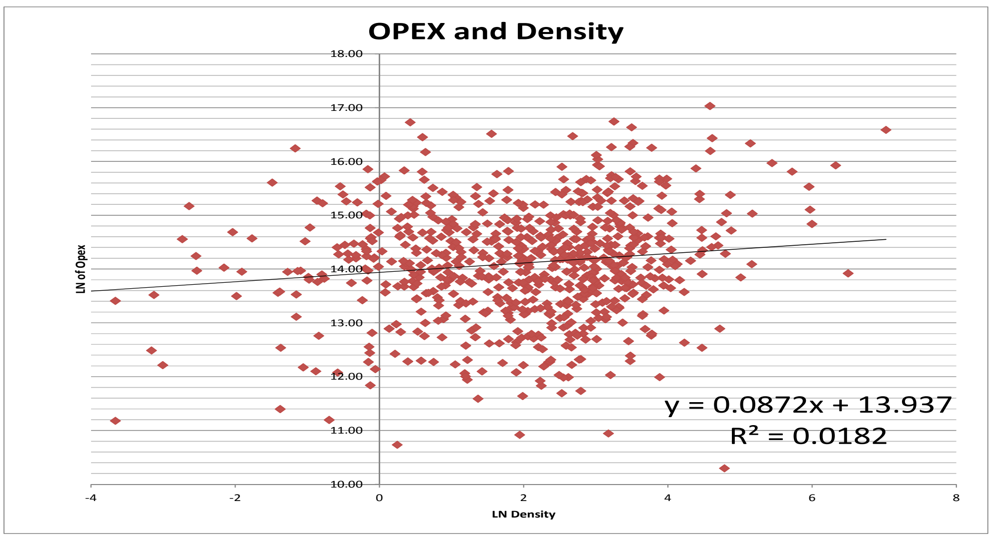

Alternate COST and Number of Study Areas in State
LN of Alt COST
LN Number of Study Areas in State
y = -0.1924x + 16.491
R² = 0.0119

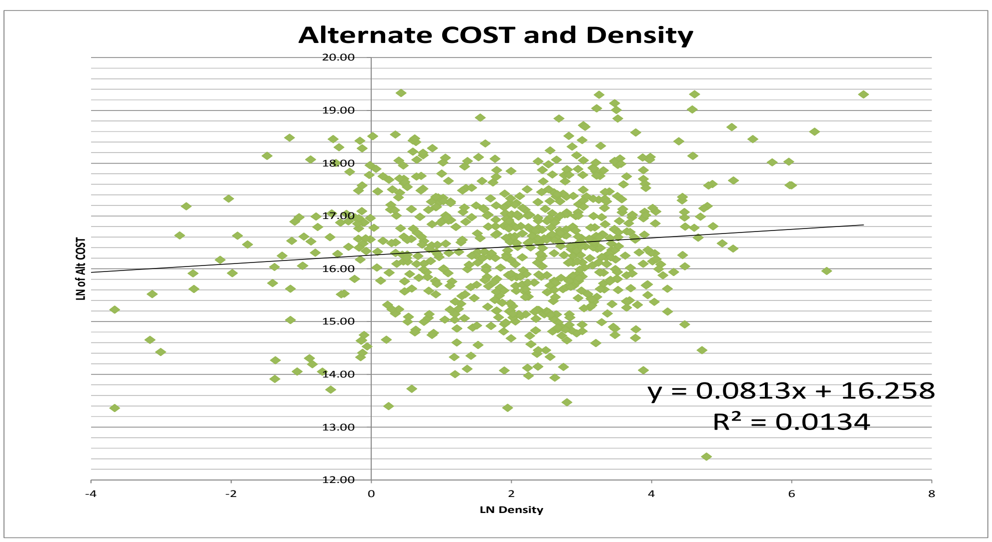

CAPEX and Pct Undepreciated Plant
LN of Capex
Pct Undepreciated Plant
y = 0.0275x + 12.847
R² = 0.1023

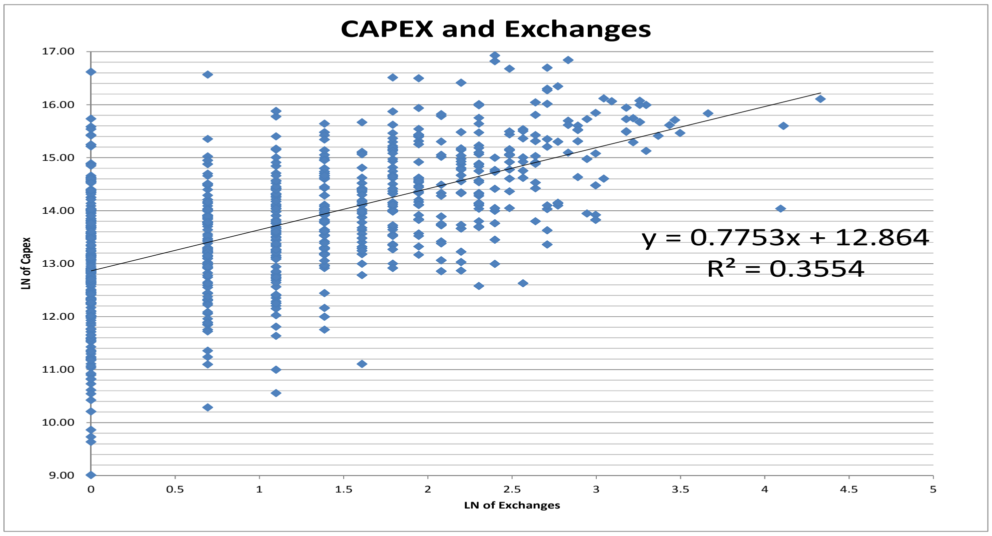

OPEX and Pct Undepreciated Plant
LN of Opex
Pct Undepreciated Plant
y = 0.0029x + 14.015
R² = 0.0017

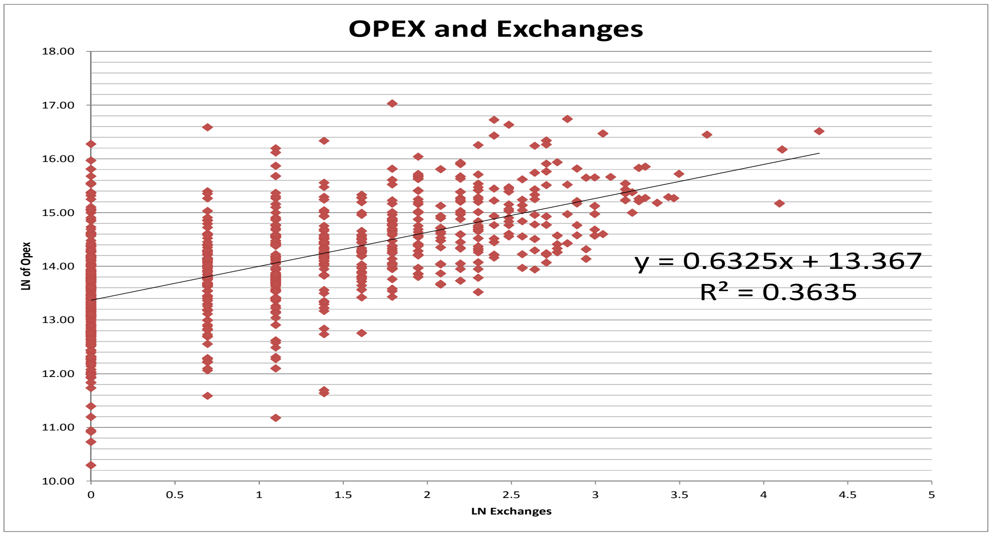

Alternate COST and Pct Undepreciated Plant
LN of Alt COST
Pct Undepreciated Plant
y = 0.0078x + 16.158
R² = 0.0107

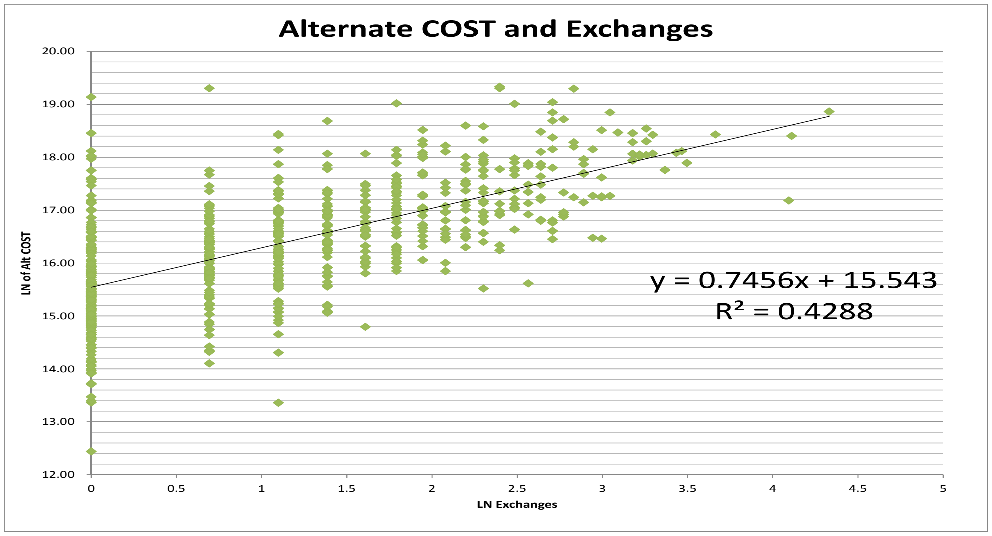

CAPEX and Density
LN of Capex
LN of Density
y = 0.0754x + 13.626
R² = 0.0088

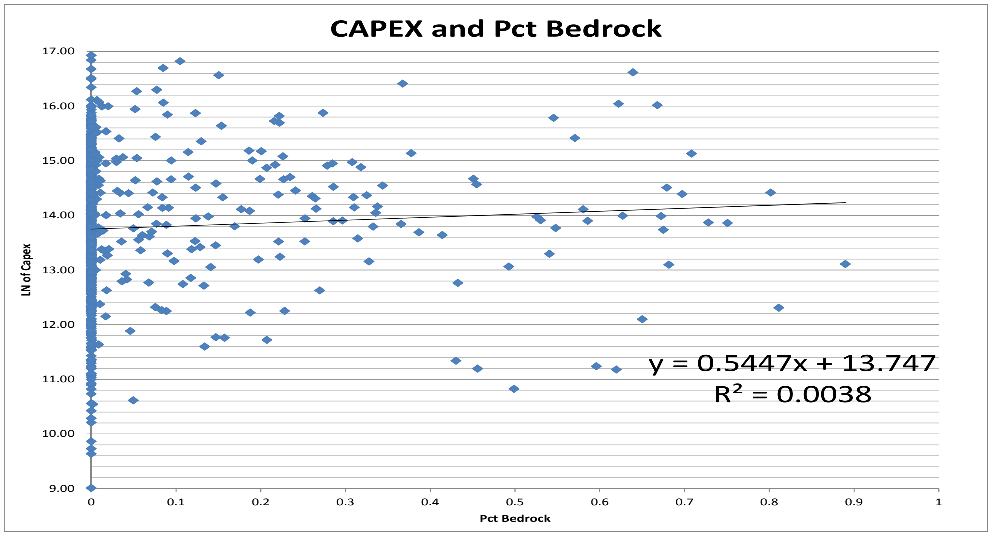

OPEX and Density
LN of Opex
LN Density
y = 0.0872x + 13.937
R² = 0.0182

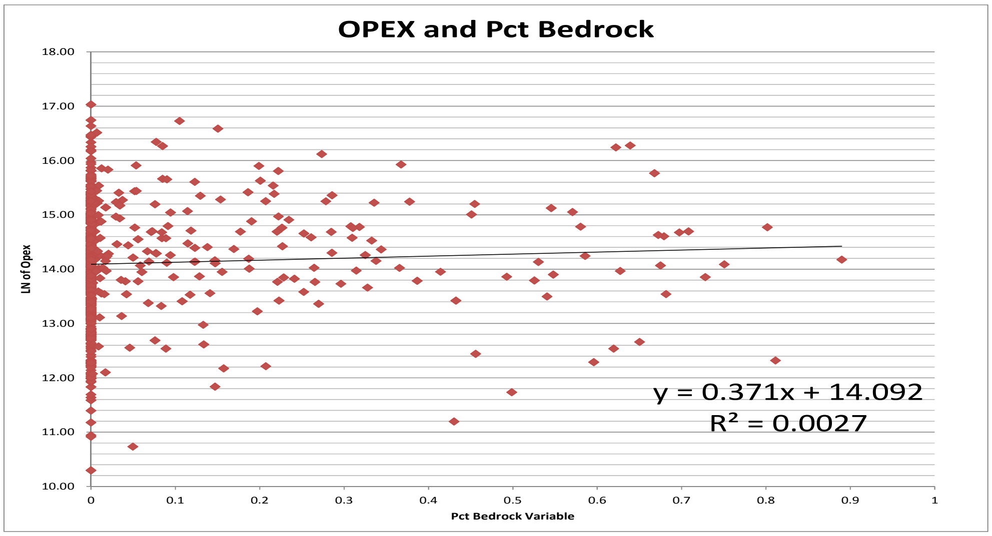

Alternate COST and Density
LN of Alt COST
LN Density
y = 0.0813x + 16.258
R² = 0.0134

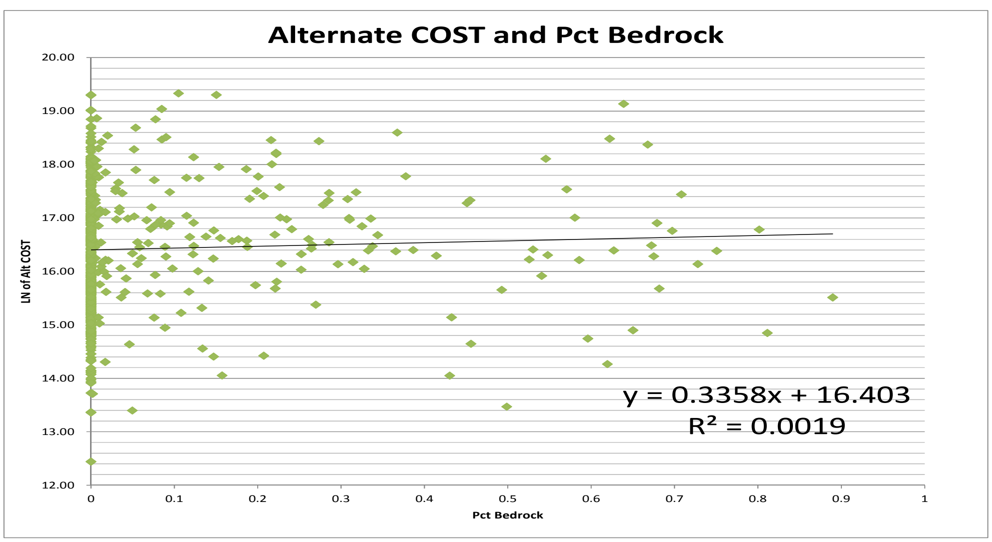

CAPEX and Exchanges
y = 0.7753x + 12.864
R² = 0.3554
LN of Capex
LN of Exchanges

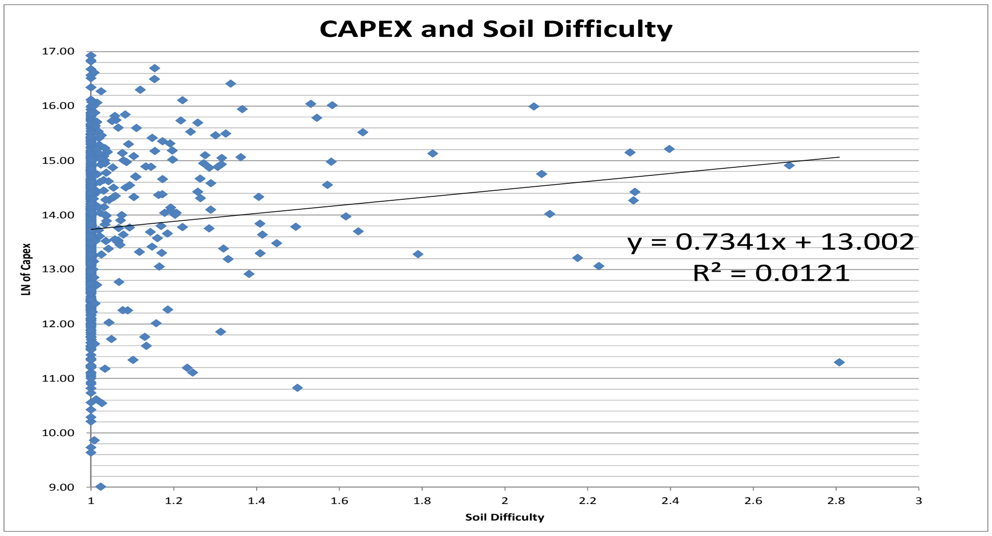

OPEX and Exchanges
y = 0.6325x + 13.367
R² = 0.3635
LN of Opex
LN Exchanges

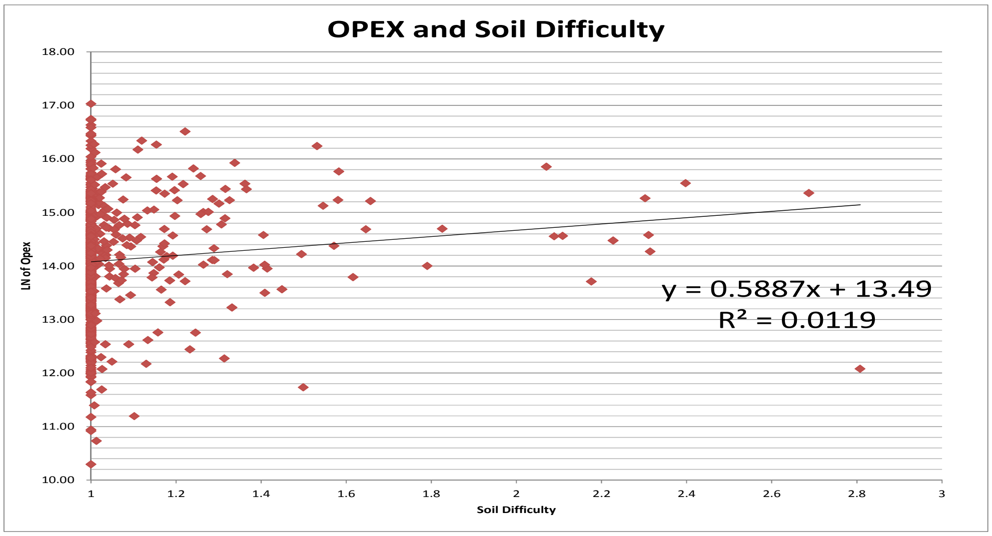

Alternate COST and Exchanges
y = 0.7456x + 15.543
R² = 0.4288
LN of Alt COST
LN Exchanges

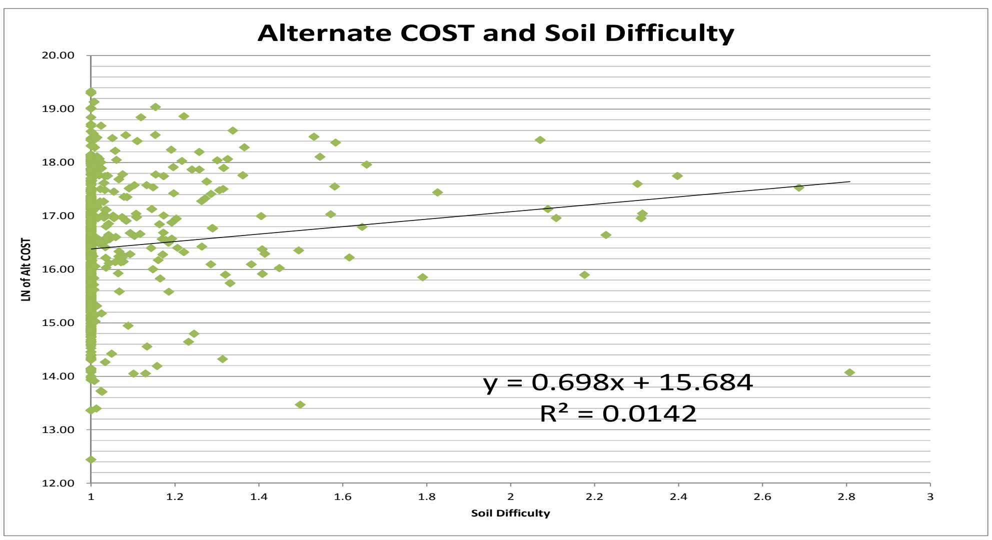

CAPEX and Pct Bedrock
LN of Capex
Pct Bedrock
y = 0.5447x + 13.747
R² = 0.0038

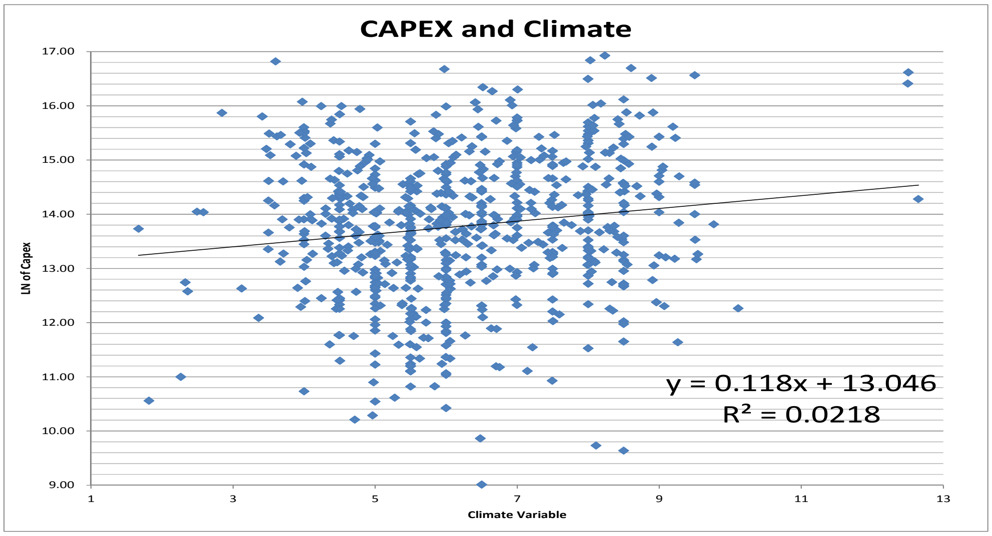

OPEX and Pct Bedrock
LN of Opex
Pct Bedrock Variable
y = 0.371x + 14.092
R² = 0.0027

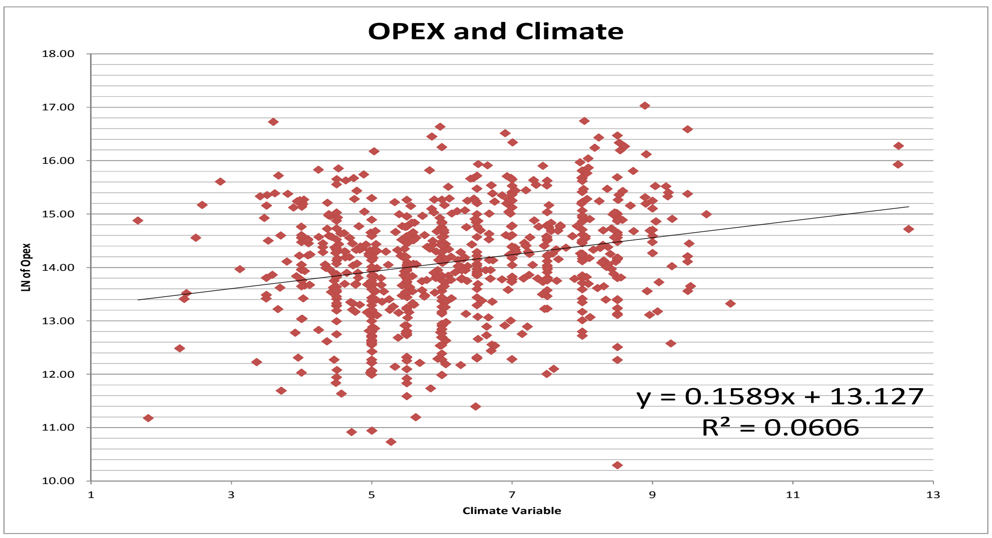

Alternate COST and Pct Bedrock
LN of Alt COST
Pct Bedrock
y = 0.3358x + 16.403
R² = 0.0019

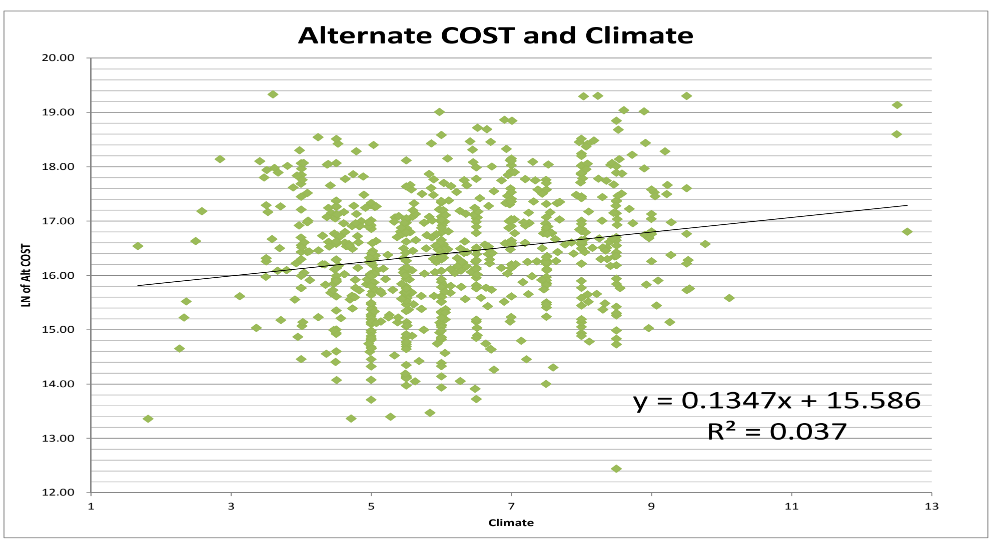

CAPEX and Soil Difficulty
y = 0.7341x + 13.002
R² = 0.0121
LN of Capex
Soil Difficulty

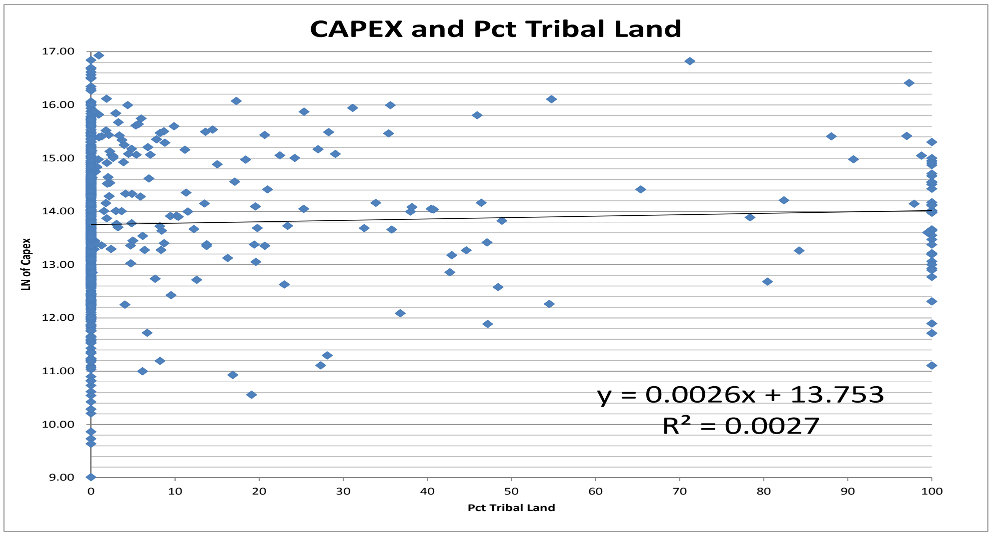

OPEX and Soil Difficulty
y = 0.5887x + 13.49
R² = 0.0119
LN of Opex
Soil Difficulty

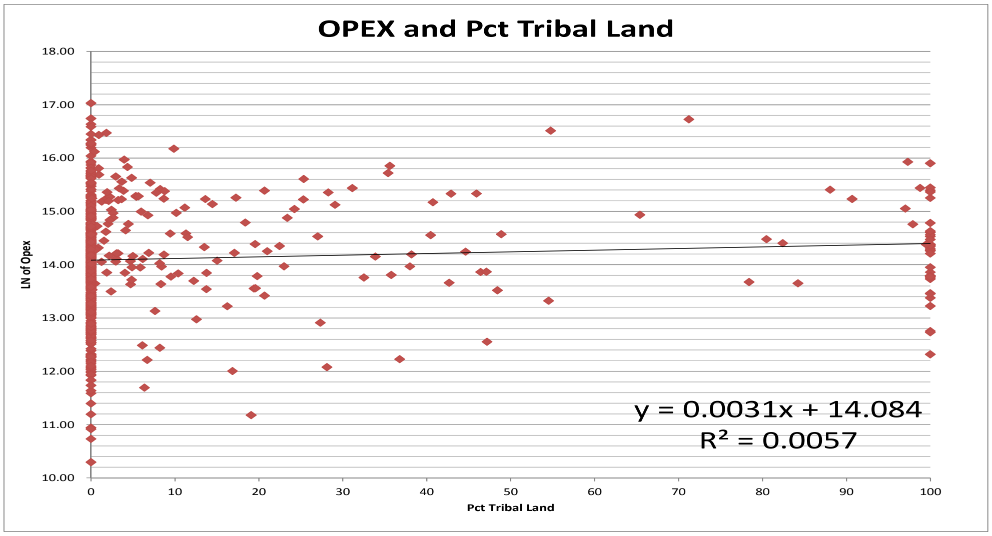

Alternate COST and Soil Difficulty
y = 0.698x + 15.684
R² = 0.0142
LN of Alt COST
Soil Difficulty

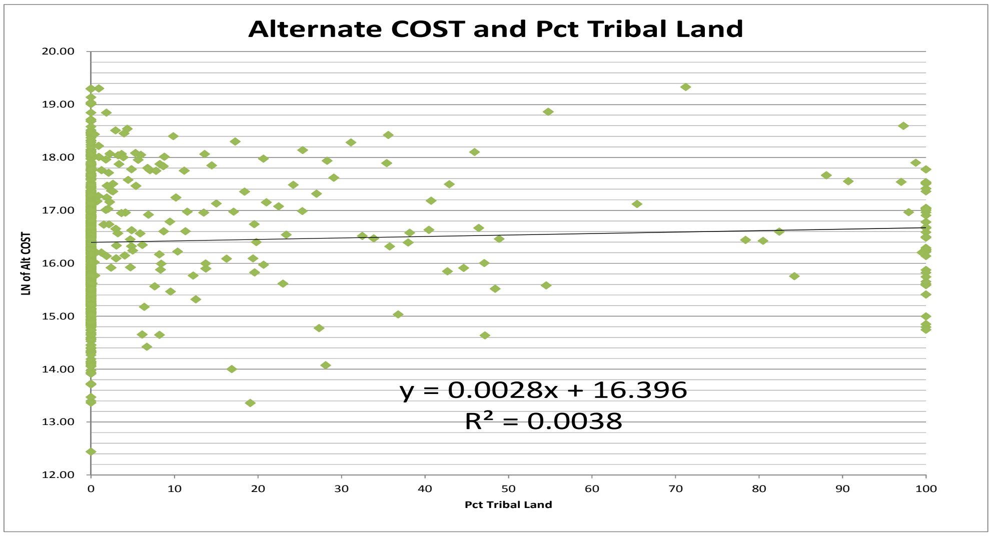

CAPEX and Climate
LN of Capex
Climate Variable
y = 0.118x + 13.046
R² = 0.0218

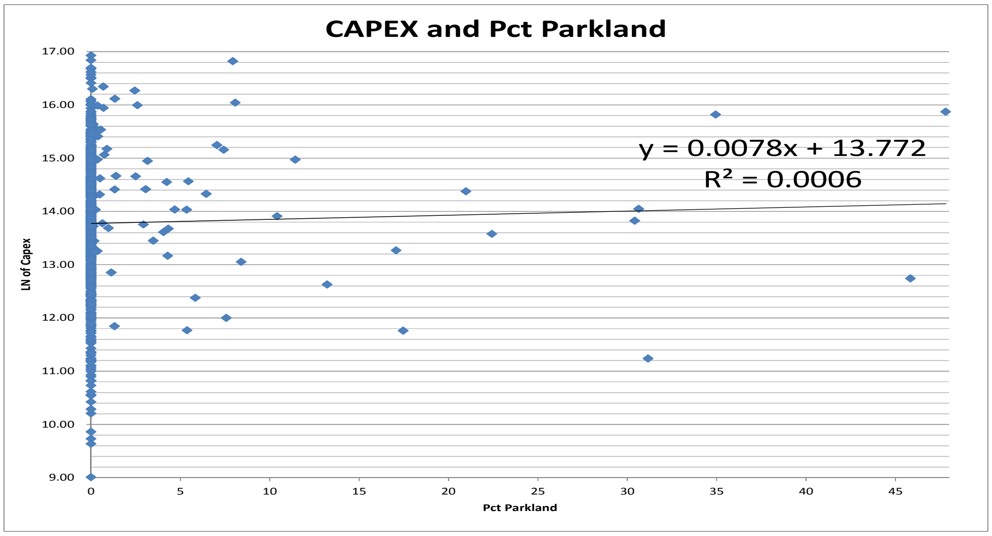

OPEX and Climate
LN of Opex
Climate Variable
y = 0.1589x + 13.127
R² = 0.0606

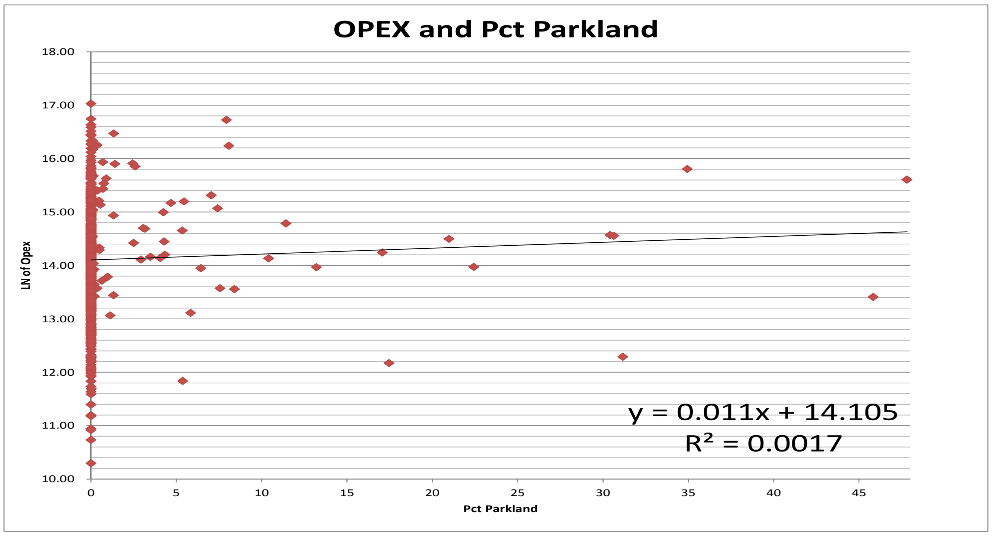

Alternate COST and Climate
LN of Alt COST
Climate
y = 0.1347x + 15.586
R² = 0.037

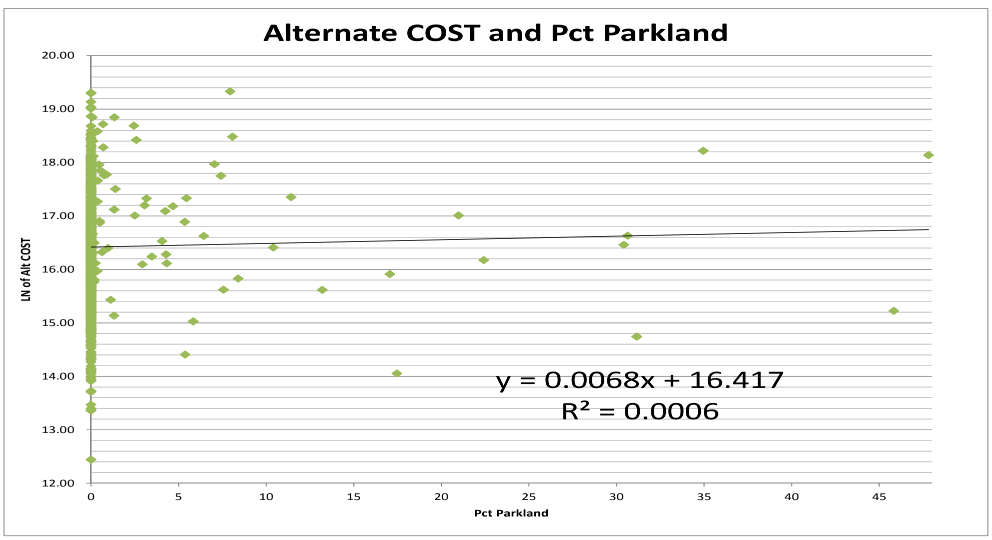

CAPEX and Pct Tribal Land
LN of Capex
Pct Tribal Land
y = 0.0026x + 13.753
R² = 0.0027

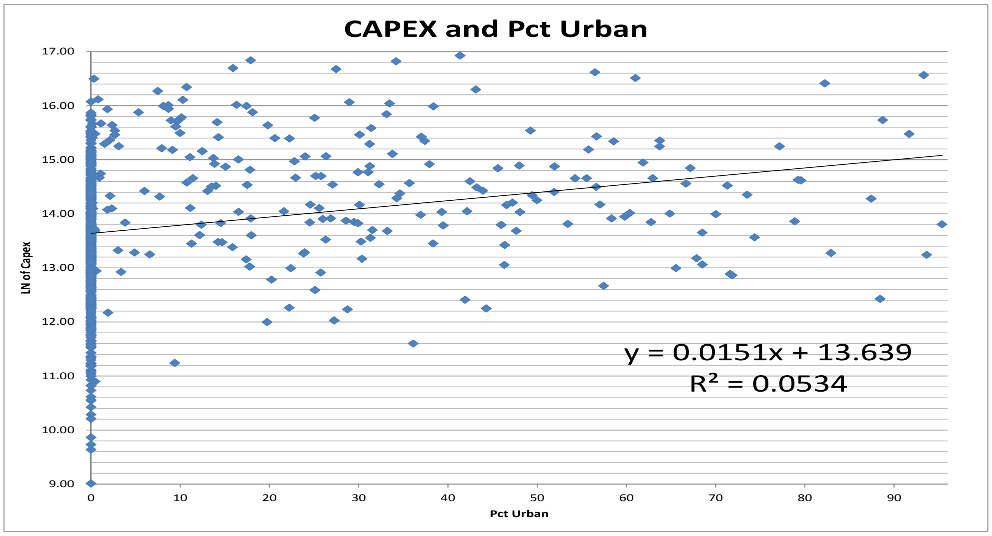

OPEX and Pct Tribal Land
LN of Opex
Pct Tribal Land
y = 0.0031x + 14.084
R² = 0.0057

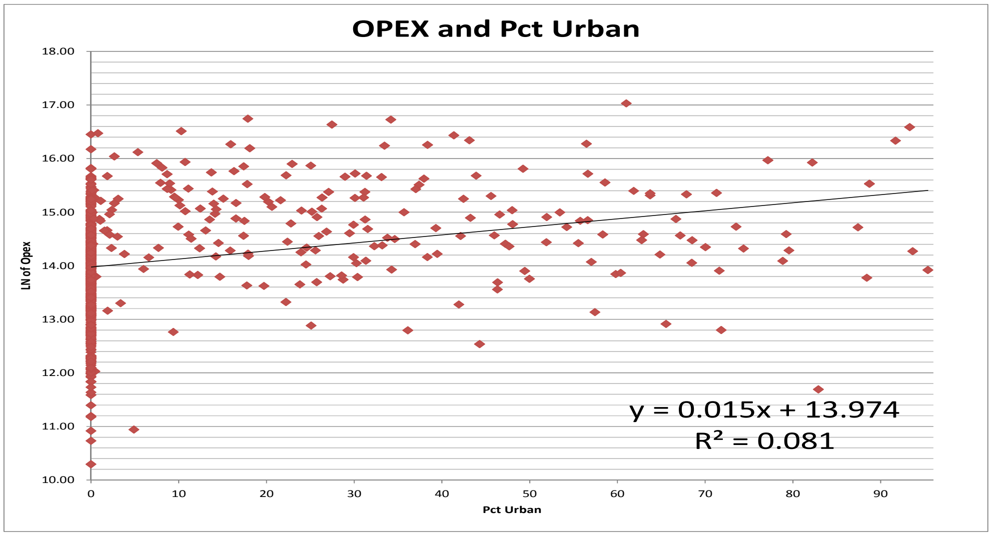

Alternate COST and Pct Tribal Land
LN of Alt COST
Pct Tribal Land
y = 0.0028x + 16.396
R² = 0.0038

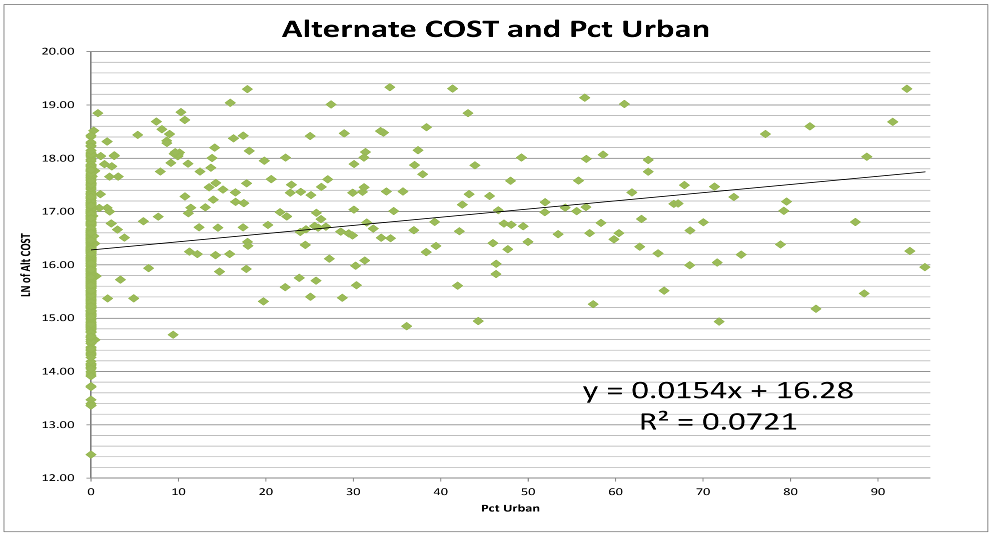

CAPEX and Pct Parkland
y = 0.0078x + 13.772
R² = 0.0006
LN of Capex
Pct Parkland

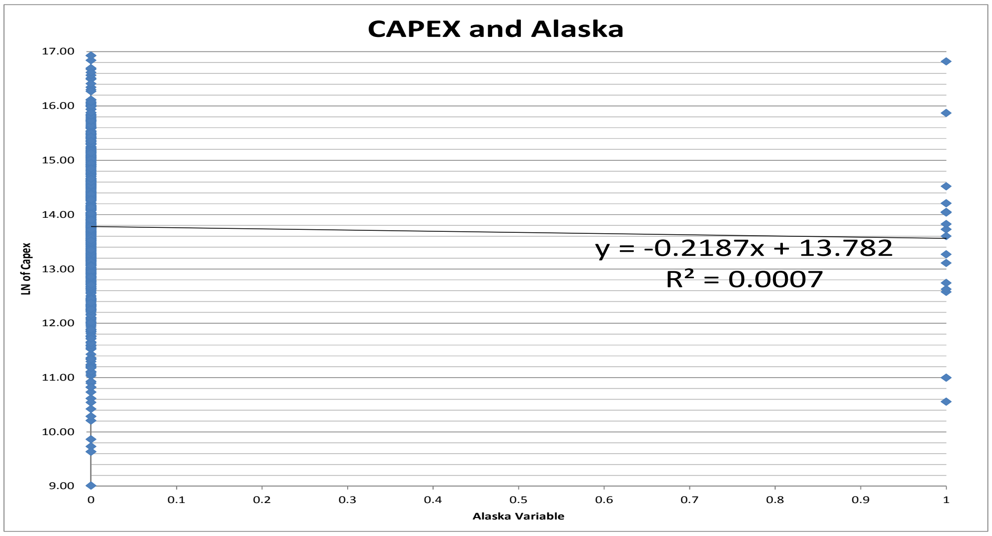

OPEX and Pct Parkland
y = 0.011x + 14.105
R² = 0.0017
LN of Opex
Pct Parkland

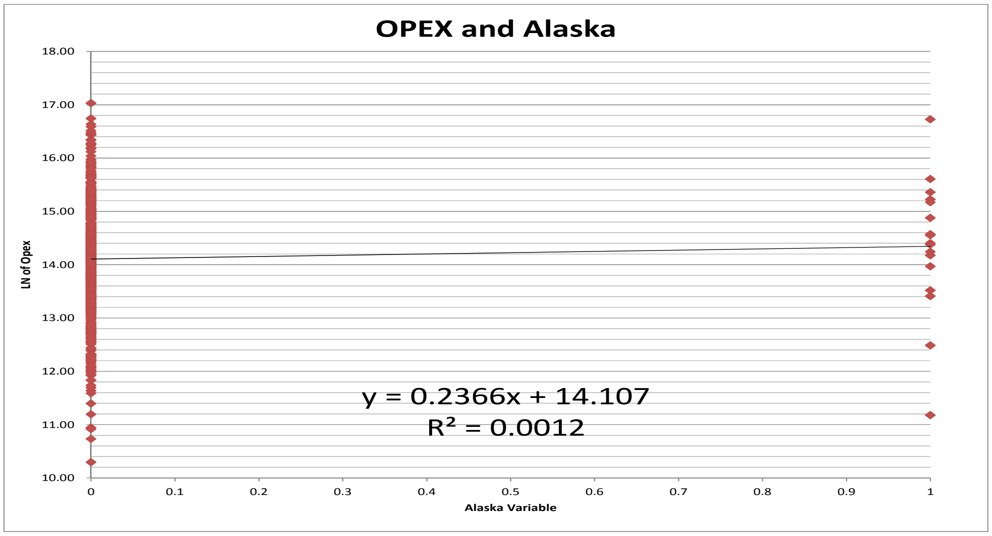

Alternate COST and Pct Parkland
y = 0.0068x + 16.417
R² = 0.0006
LN of Alt COST
Pct Parkland

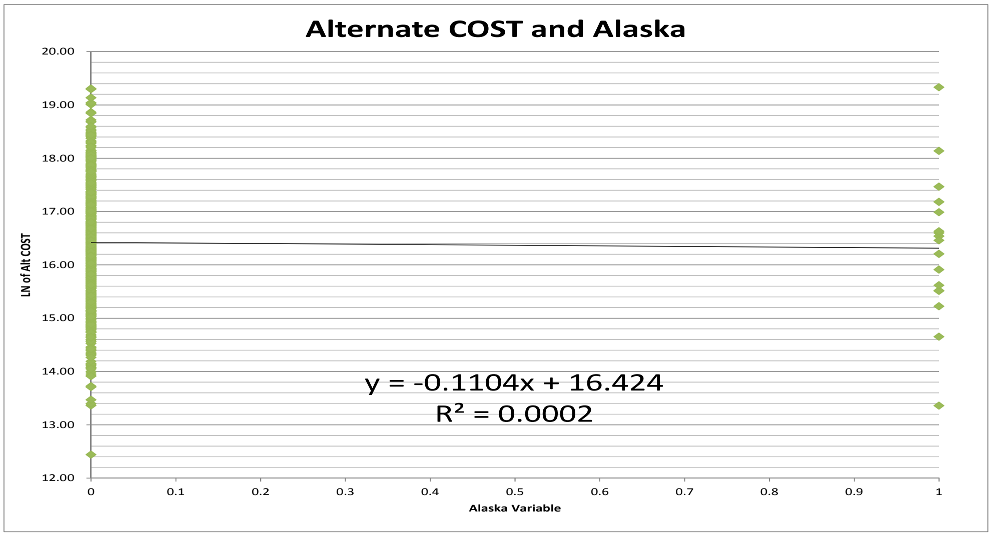

CAPEX and Pct Urban
LN of Capex
Pct Urban
y = 0.0151x + 13.639
R² = 0.0534

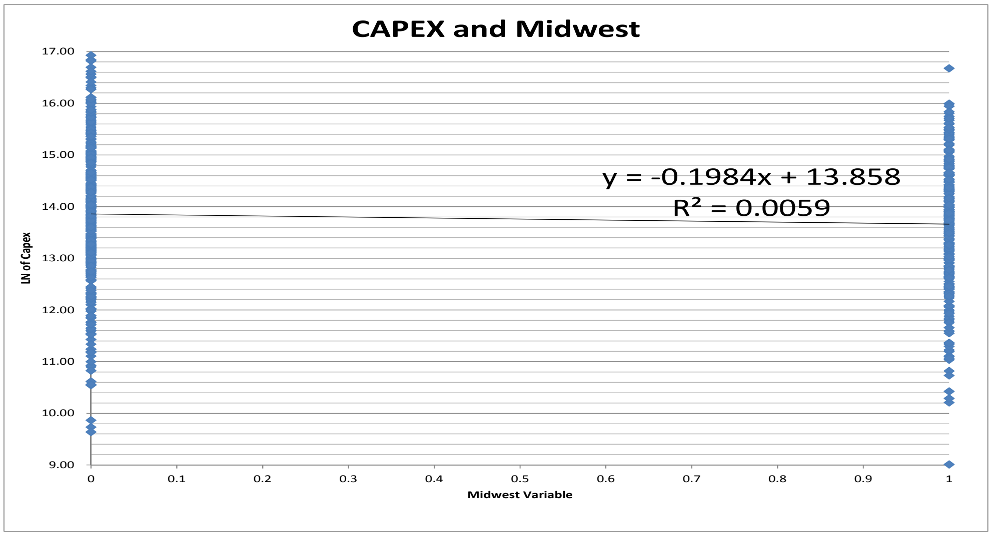

OPEX and Pct Urban
LN of Opex
Pct Urban
y = 0.015x + 13.974
R² = 0.081

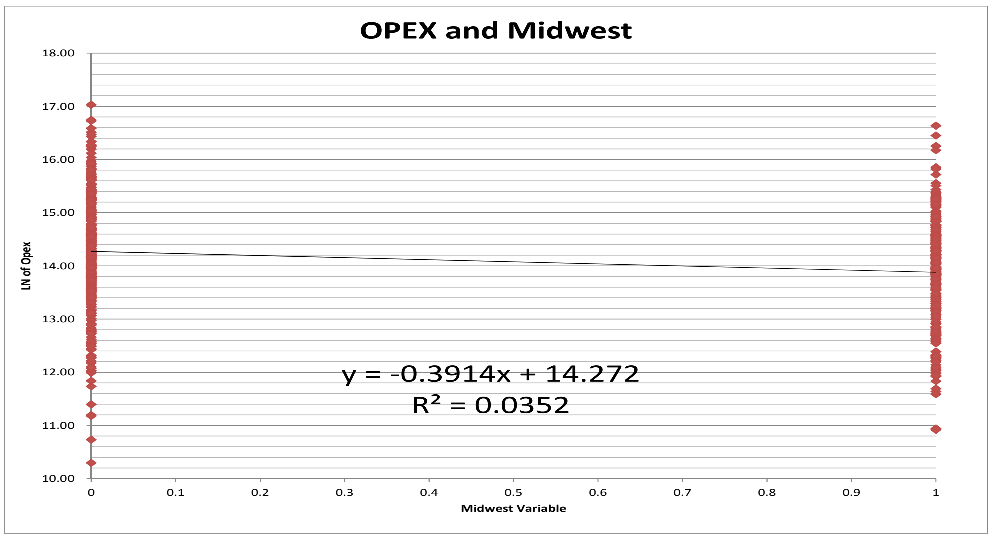

Alternate COST and Pct Urban
LN of Alt COST
Pct Urban
y = 0.0154x + 16.28
R² = 0.0721

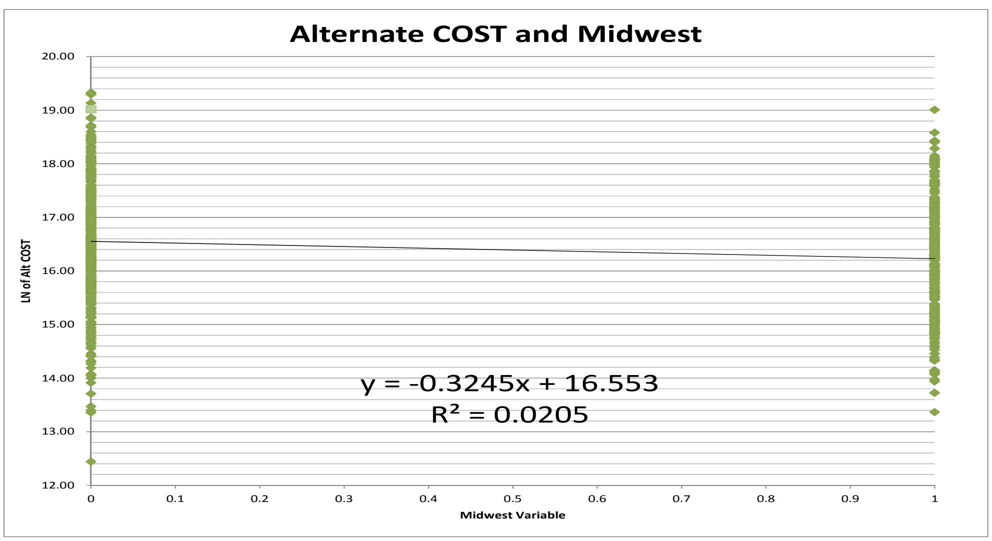

CAPEX and Alaska
LN of Capex
17.00
16.00
15.00
14.00
13.00
12.00
11.00
10.00
9.00
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Alaska Variable
y = -0.2187x + 13.782
R² = 0.0007

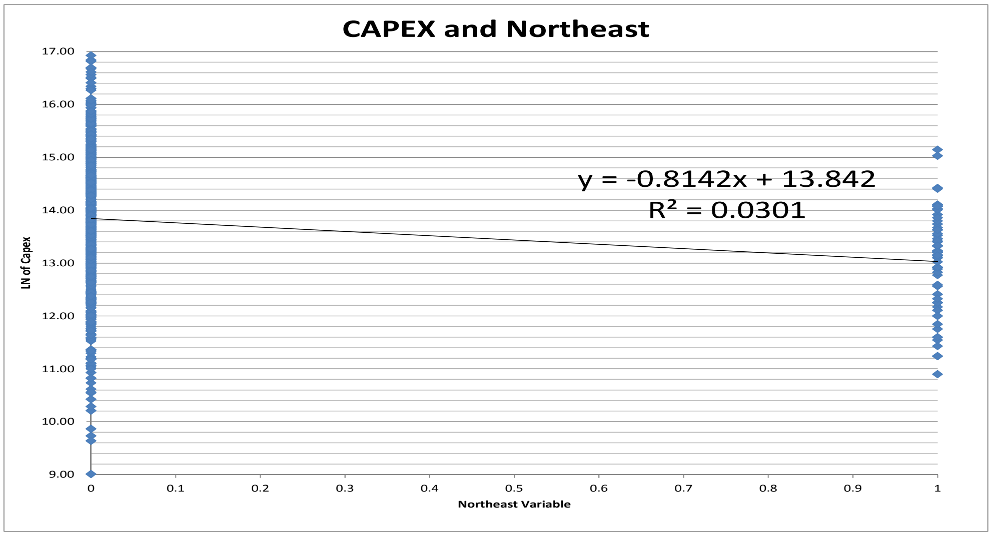

OPEX and Alaska
LN of Opex
18.00
17.00
16.00
15.00
14.00
13.00
12.00
11.00
10.00
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Alaska Variable
y = 0.2366x + 14.107
R² = 0.0012

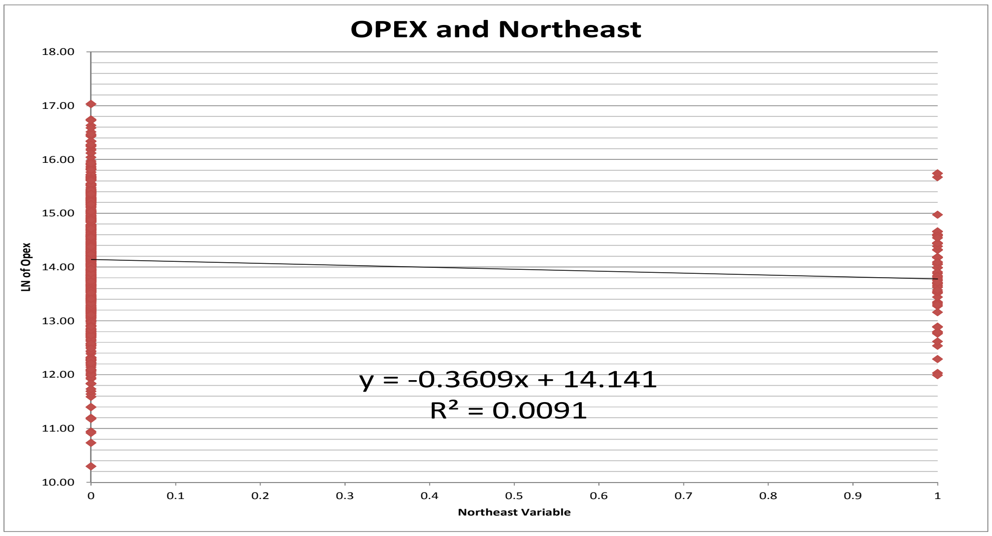

Alternate COST and Alaska
LN of Alt COST
20.00
19.00
18.00
17.00
16.00
15.00
14.00
13.00
12.00
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Alaska Variable
y = -0.1104x + 16.424
R² = 0.0002

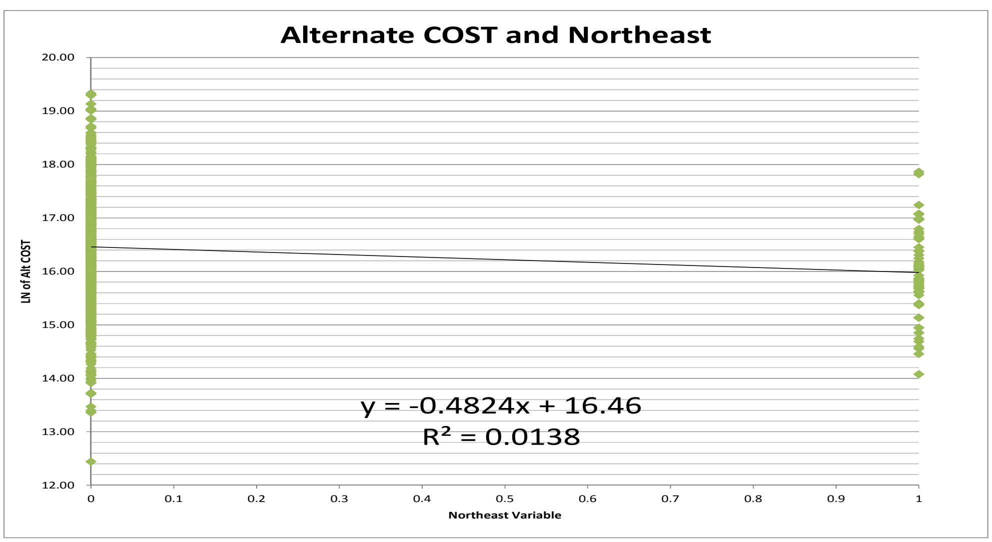

CAPEX and Midwest
y = -0.1984x + 13.858
R² = 0.0059
LN of Capex
Midwest Variable


OPEX and Midwest
y = -0.3914x + 14.272
R² = 0.0352
LN of Opex
Midwest Variable


Alternate COST and Midwest
y = -0.3245x + 16.553
R² = 0.0205
LN of Alt COST
Midwest Variable


CAPEX and Northeast
y = -0.8142x + 13.842
R² = 0.0301
LN of Capex
Northeast Variable


OPEX and Northeast
y = -0.3609x + 14.141
R² = 0.0091
LN of Opex
Northeast Variable


Alternate COST and Northeast
y = -0.4824x + 16.46
R² = 0.0138
LN of Alt COST
Northeast Variable

Appendix E: Response to Certain FCC Commentaries about the QRA

The FCC has adopted the QRA as a tool presumptively to disallow certain expenses incurred by carriers. This Appendix provides a more ample presentation of the FCC's commentaries about the value of the QRA, as outlined in the USF/ICC Transformation Order and the HCLS Benchmark Order. This Appendix also provides initial responsive comments to those commentaries.

Purpose of the QRA, USF/ICC Transformation Order, 220

> *Under our new rule, we will place limits on the HCLS provided to carriers whose costs are significantly higher than other companies that are similarly situated, and support will be redistributed to those carriers whose unseparated loop cost is not limited by operation of the benchmark methodology.*

Comment

The FCC has authorized (i) limiting HCLS funding, (ii) after determining that the costs are significantly higher than those of (iii) companies that are similarly situated, and then (iv) redistributing those funds to other companies. The FCC's intent is to limit overspending by certain carriers and to provide more funding to other carriers.

Our observations sharpen those comments, as the QRA apparently limits cost-recovery of certain carriers, *based on the assumption that the QRA can correctly determine "overspending."* Importantly, the overspending may or may not be real. Second, the assertion that the comparisons are based on "companies that are similarly situated" is only broadly true in the sense that rural carriers in all parts of the country are compared with all rural companies in all other parts of the country. The QRA compares cost companies with average schedule companies, very small companies with relatively larger companies, companies in Alaska with those in Florida, copper-based companies with those that are deploying more extensive fiber operations, among other factors. The FCC explains its use of the term "similarly situated" by referring to use of coefficients generated by the QRA.[64] Significantly, the QRA tool is used to effect the comparison *based on another assumption* that those "similarly-situated companies" can be compared by employing a single, untested statistical model, in spite of the model's R^2 which is below 68% (and inability to "explain" 30%+ of the modeled costs). Finally, the reallocation of funding to other carriers reflects another assumption that the reallocated funding achieves the purpose of universal service funding better in those cases than it would if it were allocated to certain reported higher-cost areas.

We admit readily that we do not know whether the funding should be disallowed in certain cases. However, we are clear that there are fundamental, untested assumptions in the FCC's use of the QRA. That is, the FCC assumes that the QRA correctly identifies "overspending," that the model is a sufficient and effective "comparison" without the intervention of human judgment, and that the limited HCLS funding is better spent in other regions than in the regions that are disallowed. We respectfully suggest

---

[64] USF/ICC Transformation Order, Appendix H, Footnote 1: "The term 'similarly-situated peers' means that, based on data from all the carriers in the analysis, if there were (hypothetically) 100 study areas with independent variable values that were nearly the same as those with the study area in question, 90 of them would be expected to have values equal to or less than the 90th percentile prediction. It does not mean the carriers with the most similar number of loops (or values of the other variables)."

that those assumptions deserve to be better tested to avoid harm to customers and the statutory public policy.

Predictability of the QRA, USF/ICC Transformation Order, 220

> *We note that the fact that an individual company will not know how the benchmark affects its support levels until after investments are made is no different from the current operation of high-cost loop support, in which a carrier receives support based on where its own cost per loop falls relative to a national average that changes from year to year. Even today, companies can only estimate whether their expenditures will be reimbursed through HCLS. In contrast to the current situation, the new rule will discourage companies from over-spending relative to their peers. The new rule will provide additional support to those companies that are otherwise at risk of losing HCLS altogether, and would not otherwise be well-positioned to further advance broadband deployment.*

Comment

The FCC acknowledges the "unpredictability" associated with the use of a model, and correctly highlights that there was unpredictability in the former regime. However, it appears to us that there is no comparison between the *degree of unpredictability* in the new approach and the relatively lesser uncertainty in the previous capped high cost fund.

Prior to 2012, a cap on HCLS resulted in allocations of cost recovery based on a National Average Cap per Loop ("NACPL"). From the time when the NACPL was capped in 2001, the effective NACPL rose virtually every year by an average of $21 per loop. So a carrier that planned for capital investment could review its proposed expenditures and assume that it would receive recovery if it were within certain ranges as the NACPL movement was relatively regular. If the carrier were a duck hunter, it would simply understand that it needed to "lead" the duck in flight to hit its target (cost recovery based on the NACPL) in the early morning light, assuming good weather conditions. The QRA is much, much less predictable because, in the new regression analyses, there are far more variables that have been introduced, and they change in an annual recalculation of the model. The new QRA is duck hunting when the winds are high, the distance is farther, and, for sport, there is no light. Both involve unpredictability, but the latter case introduces many complex variables that will necessarily affect a carrier's confidence in calculating investment in high-cost areas. This last point is most important. A carrier may be relatively confident if its costs are not high (because its service region is not terribly expensive or its plant was upgraded many years ago), but if the service region is inherently high cost or the deployment requires higher investment in new technologies, the recovery mechanism is not at all predictable. In our discussions with carriers, we are clear that the QRA is most likely to result in lower investment in higher-cost regions because of the unpredictability of the QRA. The QRA is unpredictable for the highest-cost areas, and will have an effect that insidiously undercuts a major purpose of USF. As noted earlier, the 2013 revision of the QRA has resulted in approximately 60% more carriers that have reductions compared with the 100 carriers affected in 2012.

VALUE AND CHARACTERISTICS OF QRA, USF/ICC TRANSFORMATION ORDER, APPENDIX H, 6-7

We considered using an ordinary least squares-based ["OLS"] analysis to set the caps, but decided that quantile regression was preferable for two reasons. First, error terms in bivariate OLS models of each algorithm step on the loops variable exhibit heteroscedasticity. While ordinary least squares-based analyses such as weighted least squares can certainly deal with heteroscedasticity, it complicates efforts to deal with other problems such as outliers and non-Gaussian error terms. Further, ordinary least squares can produce biased parameter estimates in the presence of outliers. Ordinary least squares has methods available for dealing with outliers, such as excluding them from the analysis or using dummy variables to deal with them, but that requires exercise of judgment as to which observations are truly outliers. Also, given the data currently available to the Commission, distinguishing between study areas with high idiosyncratic costs (i.e., those that truly are the most expensive-to-serve areas) and others with excessively high cost (e.g., due to imprudent or unnecessarily large past investments) is challenging. Further complicating matters, some carriers may enjoy especially low costs compared to their peers for idiosyncratic reasons. While these observations would be outliers, they would be masked by the virtue that they are somewhat "too low" and therefore it would be difficult to properly identify and deal with those outliers. Thus, simply looking only for observations that are too high may be insufficient. When using ordinary least squares, failing to account for all outliers (including the difficult-to-find outliers that are "too low") could bias the regression coefficients which would then bias payments to carriers. Quantile regression solves this problem.

Comment

The FCC is using a quantile regression analysis rather than a least squares regression because of, at least, two positive characteristics in the QRA. First, there is heteroscedasticity. Its opposite—homoscedasticity—refers to a regular and even, tightly-grouped spread on a scatterplot depicting independent and dependent variables. By contrast, heteroscedasticity refers to a situation where the spread may be tight in one part of the scatterplot and less predictable in another. So, for example, costs might be predicted more easily when a carrier serves a larger number of loops but less accurately when it serves fewer loops. A statistician would prefer homoscedasticity so that the predictive value is good across the entire range. Second, the FCC also highlights that "outliers"—very high or very low values—can bias the result by making averages too high or too low. According to the FCC, quantile regression "solves" this problem.

We have used regressions and have reviewed the analyses of statisticians who use these tools. Notably, the FCC cited the work of Dr. Roger Koenker who, together with another statistician, developed the original concept of the quantile regression analysis.[65] Dr. Koenker prepared a document that was filed as part of a larger response to the USF/ICC Transformation Order.[66] Dr. Koenker's comments were relatively narrow in focusing on the statistical model, rather than on the content or data input issues which are raised in other parts of this White Paper. He raised substantive statistical concerns about (i) the use of

[65] USF/ICC Transformation Order, Appendix H, para. 8.

[66] National Exchange Carrier Association et al. Comments, WC Docket No. 10-90 et al., at App. E, 1 (filed Jan. 18, 2012) (Roger Koenker, "Assessment of Quantile Regression Methods for Estimation of Reimbursable Cost Limits"); ("Koenker Appendix E").

two dependent variables (capex and opex) because high costs related to capex or opex might be offset by the low costs for the other, and he recommends combining the two into one dependent variable, again, for statistical reasons; (ii) the introduction of "insignificant" variables, which the FCC claims does not bias the model, but in Dr. Koenker's estimation does "lead to deterioration in the performance of the predictions"; and (iii) the FCC's transformation of certain log values where it substitutes the number 1 in many instances where log(0) occurs; Dr. Koenker refers to this as "cavalier treatment of the log transformation" which he notes could have the effect of lowering the value created in regressions. He goes on to conclude that "[m]ore serious problems arise for both mean regression and quantile regressions when such perturbations introduce outliers in the explanatory variables. Such so-called leverage points can be highly influential . . . and can do serious damage to the fitted model."[67]

Our point here is not to engage in a theoretical discussion regarding the value of one or the other statistical model. However, the FCC appears to rely on a model that it argues is effective in predicting appropriate costs and in identifying excessive expenditures. The statistician to which the FCC points raises important questions about the way in which this model is used—in predicting two dependent variables, in introducing insignificant "covariates" or variables, and in using certain methods such as log transformations that could bias the outcomes. We believe that there are broader questions about the advisability of using such a presumptive model, but we note that even the statistical methodology is disputed. When combined with insufficient data sources and relatively low predictive value, the question must be raised about the sufficiency of such a presumptive tool.

The FCC's scope in implementing the methodology, HCLS Benchmark Order, 41

> *We also reject the argument that implementing these benchmarks will undermine the predictability or sufficiency of support. At the outset, we note that this general argument effectively seeks reconsideration of the Commission's policy judgment to adopt a rule imposing limits on capex and opex in the first instance, which is beyond the scope of this order to implement a methodology as directed by the Commission.*

Comment

The FCC is implementing the USF/ICC Transformation Order about which it requested additional comments. The purpose, if we have it correct, was that the FCC seeks to limit overspending and assure that sufficient and predictable funds are available to meet policy goals related to universal service.

The FCC has modified the methodology related to the QRA from the initially-proposed approach in the USF/ICC Transformation Order to the formulation in the HCLS Benchmark Order to accomplish the purpose more effectively. It is now rumored that further changes are being considered. While the authors are not attorneys, we believe it is possible to accomplish the FCC's purpose in limiting excessive spending and in assuring that other carriers receive sufficient funding in further modifications. Those changes include assuring that data sources are accurate and ample, that appropriate cost causation is tested, and that there are potential reviews of the predictive value of the QRA. Specifically, a preliminary annual or biennial review of the high-cost results to determine whether they are in fact signals of excessive spending might be more appropriate. Such an approach would better identify whether the

---

[67] Koenker Appendix E, pp. 1-8, quotations from pp. 7-8; see USF/ICC Transformation Order, Appendix H, para. 22 for the discussion of the FCC's log transformations.

spending is excessive, whether the model can be improved, and whether the funding can be better allocated when it is definitively assigned.

Appendix F: Author biographies

Vincent H. Wiemer, Principal

Vincent H. Wiemer is a Principal and founder at Alexicon Consulting, a management consulting firm that provides financial, regulatory, and advisory services to the independent telecommunications industry. Mr. Wiemer's practice concentrates on financial modeling, strategic planning, regulatory impact analysis, and business development for his clients. He is a popular industry speaker and has presented such diverse topics as metrics, effective communications, incentives, and personal accountability among others. Prior to working in the telecommunications industry, Mr. Wiemer provided public accounting and consulting services to a spectrum of industries including energy providers, government agencies, and major hotel chains. Mr. Wiemer has a bachelor's degree in business administration from the University of Tulsa and earned his Certified Public Accountant license in Oklahoma.

Michael J. Balhoff, CFA, Senior Partner

Michael J. Balhoff, CFA, is a Senior Partner and co-founder at Charlesmead Advisors, LLC, and is Managing Partner at Balhoff & Williams, LLC, a professional services firm that provides financial-regulatory consulting and advisory services to companies, investors and policymakers in the communications and energy industries. Charlesmead Advisors is an investment banking firm that provides financial advisory services, including valuation as well as merger and acquisition services to telecommunications companies. Before founding the Charlesmead and the predecessor firm to Balhoff & Williams, Mr. Balhoff headed the Technology and Telecommunications Equity Research Group at Legg Mason and, in the final seven of his sixteen years as a senior analyst at Legg Mason, he covered equities in the incumbent local exchange carrier industry. Prior to joining Legg Mason in 1989, Mr. Balhoff taught at the graduate and undergraduate levels. Mr. Balhoff has a doctorate in Canon Law and four master's degrees, including an MBA—concentration finance—from the University of Maryland. He is a CFA charterholder and is a member of the Baltimore Security Analysts Society. Mr. Balhoff has been named in six annual awards as a Wall Street Journal All-Star Analyst for his recommendations on the Telecommunications industry. His coverage of telecommunications, and especially rural telecommunications, was named by Institutional Investor magazine as the top telecommunications boutique in the country in 2003.