

AD-A033 973

CALIFORNIA UNIV LOS ANGELES WESTERN MANAGEMENT SCIE--ETC F/G 5/3
CUSTOMER AGGREGATION IN DISTRIBUTION MODELING (U)

OCT 76 A M GEOFFRION

N00014-75-C-0570

UNCLASSIFIED

WMSI WORKING PAPER-259

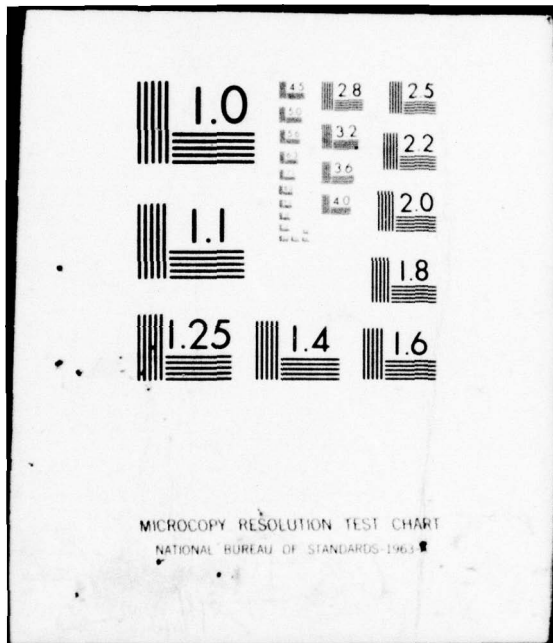
NL

1 OF 1
AD
A033973



END

DATE
FILMED
2-77



WESTERN MANAGEMENT SCIENCE INSTITUTE ✓
University of California, Los Angeles

C
B.S.

ADA 033973

14
WMSI Working Paper 259

6
CUSTOMER AGGREGATION IN DISTRIBUTION MODELING

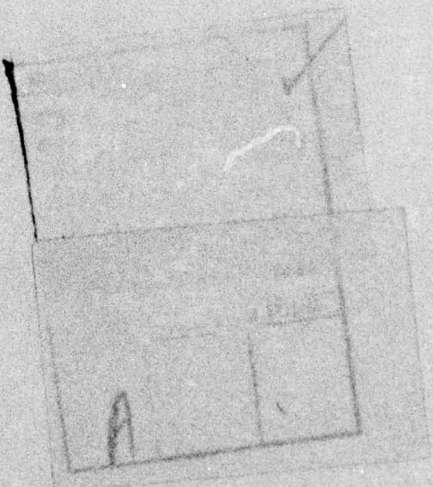
10
ARTHUR M. GEOFFRION
Graduate School of Management

11
October 1976

12
19p

15
Contract No. N00014-75-C-0570

DDC
NOV 23 1976
REGISTERED



This paper was partially supported by the National Science Foundation and by the Office of Naval Research.

N00014-75-C-0570 ✓

072 320
bpg

ABSTRACT

↙ The issue of customer aggregation arises frequently in modeling physical distribution systems. We ~~address~~ ^{is addressed} this issue for a class of single commodity models that includes the classical transportation and capacitated facility location problems as special cases. For any proposed aggregation of customers, an a priori upper bound is given on the amount of suboptimality thereby induced in the model. This bound is of practical use because it can provide a rigorous justification of aggregations that are plausible on the basis of geographic proximity or some other natural criterion. It also suggests a novel way of using standard clustering techniques to discover customer aggregations with small associated a priori error bounds. The analytical technique used to derive these results should prove useful for obtaining similar a priori bounds for other classes of models. ↗

RECEIVED BY	
NYC	Wife Center <input checked="" type="checkbox"/>
DC	Out Center <input type="checkbox"/>
UNIVERSITY	
DATE	
BY	<i>Stewart</i>
DISTRIBUTION/AVAILABILITY CODE	
Dist.	AVAIL. and/or SPECIAL
A	

CUSTOMER AGGREGATION IN DISTRIBUTION MODELING

Modeling is often said to be an art rather than a science [4]. One reason is that rigorous justification is seldom available for many of the design choices faced during the modeling process. This paper attempts to put a little more science into a very commonly occurring design question in the area of distribution modeling: how to aggregate the usually very large number of individual customers into a more tractable number of groups. The need for such aggregation springs from the desire to avoid excessive data development costs while building a model, or an excessive amount of computer time or main storage to solve it. It is not at all uncommon in practical studies to aggregate several thousand customers down to one or two hundred demand zones on the basis of geographical proximity and type of customer. Traditionally this has been done and defended on the basis of common sense because apparently there is no known rigorous and practically feasible approach to this task. Our aim is to remedy this deficiency.

The spirit of the present effort is akin to that of a contemporaneous paper by the author [3] on the subject of a priori error bounds for the aggregation of procurement commodities. Although the results are of a similar type, the details and analytical techniques are in fact quite different.

1. Main Results

The following model serves as the vehicle for our main results.

$$(1) \quad \text{Minimize}_{y,z} \quad \sum_{kl} d_{kl} y_{kl} + \sum_k F_k \left(\sum_l q_l y_{kl}, z_k \right)$$

$$(2) \quad \text{subject to} \quad \sum_k y_{kl} = 1, \quad \text{all } l$$

$$(3) \quad \underline{v}_k z_k \leq \sum_l q_l y_{kl} \leq \bar{v}_k z_k, \quad \text{all } k$$

(4) $0 \leq y_{kl} \leq 1$, all kl

(5) $z_k = 0, 1$ for all k and $z \in Z$.

A conventional interpretation is as follows:

k indexes the possible *facilities* from which customers can be served

l indexes the *customers*

y_{kl} a variable giving the fraction of the annual needs of customer l (for goods or services) satisfied by facility k

z_k a binary variable indicating whether facility k is selected for use

d_{kl} annual variable costs incurred if the full needs of customer l are met from facility k

$F_k(\cdot)$ other annual costs associated with facility k as a function of its annual throughput

q_l a quantity (assumed ≥ 0) measuring the annual needs of customer l

\underline{v}_k (\bar{v}_k) a lower (upper) limit on the annual throughput permissible for facility k if it is used

Z an arbitrary constraint set on z .

It is understood that a list of allowable (k,l) links is given to reflect which candidate facilities are allowed to serve which customers. All summations run only over allowable combinations.

The model as stated is a classical capacitated facility location problem with possibly nonlinear warehouse-related costs and additional constraints. No assumptions have been made regarding the form of the

functions F_k , so they could incorporate an annual fixed cost associated with the use of facility k (which is customarily expressed as $f_k z_k$) and the influence of economies or diseconomies of scale. See, e.g., [2] for a recent discussion of similar models.

The model need not necessarily involve facility location decisions. By taking Z to require $z_k = 1$ for all k , F_k to be identically 0 and v_k to be 0 for all k , (1) - (5) reduces to the classical transportation problem (with the flow variables scaled by destination demands).

What does one really mean by "aggregating" a subset L of customers?

An important type of customer aggregation is *pro rata by demand*, which amounts to introducing the following additional constraints:

(6) for each k , the y_{kl} 's must be identical over $l \in L$.

An implicit assumption, and one that we adopt henceforth, is

(7) for each $l \in L$, the same kl links exist.

An obvious consequence of (6) is that it permits variables and constraints to be eliminated. The net effect is that the mathematical structure of (1) - (5) remains unchanged, but with the number of l indices reduced everywhere by $||L|| - 1$, where $||L||$ is the number of indices in L . An illustration is presented in Sec. 2, and subsequently we shall generalize to the case where several aggregation subsets are involved simultaneously.

Aggregation clearly can be expected to result in a model with higher minimum cost. The question is how much higher. Our main result along these lines is the easily calculated *a priori* bound given in the following theorem. The proof is in the Appendix.

The notation $v[\cdot]$ stands for the optimal value of any optimization problem to which it is applied.

Theorem (Customer Aggregation). Let L be any subset of customers satisfying (7). Then

$$(8) \quad v[\text{problem (1)-(5)}] \leq v[\text{problem (1)-(6)}] \leq v[\text{problem (1)-(5)}] + \epsilon_L$$

where

$$(9) \quad \epsilon_L \triangleq \sum_{l \in L} \max_k \left\{ \frac{q_l \sum_{l' \in L} d_{kl'}}{\sum_{l' \in L} q_{l'}} - d_{kl} \right\}$$

Moreover, if (\tilde{y}, \tilde{z}) is ϵ -optimal in the aggregated problem (1)-(6), then it is feasible and $(\epsilon + \epsilon_L)$ -optimal in the original problem (1)-(5).

This fundamental result permits a modeler to determine a priori -- in advance of solving any version of the model -- an upper bound on the amount by which any proposed customer aggregation could diminish model accuracy. No longer is it necessary to rely entirely on intuition and cumbersome numerical experimentation with pilot models.

2. A Numerical Illustration

The above ideas will be illustrated numerically in miniature using a small classical transportation problem.

Consider a firm with facilities in Seattle, Los Angeles and Houston, and with customers in Dallas, Chicago, Atlanta, Pittsburgh, New York and Boston. An aggregation of the northeastern customers is desired if this can be done without introducing excessive error. The aggregation analysis will be conducted using approximate transportation costs based on available regression relationships of cost against distance. This enables transportation costs to be estimated inexpensively by computer knowing only the locational coordinates of each origin and destination. After the

aggregation analysis is completed, more accurate transportation costs would be developed for the smaller (aggregated) model and used thereafter.

Table 1 lays out the full transportation problem in a traditional format using approximate transportation cost data (taken here to be proportional to distance at the rate of one dollar per hundredweight per thousand miles). Supplies and demands are given in thousands of hundredweight. Disregard the optimal solution shown for the time being.

	DAL	CHI	ATL	PIT	NY	BOS			
SEA	2.078	2.013 (10)	2.618	2.465	2.815	2.976 (5)	15	S	
LA	1.387 (10)	2.054	2.182	2.426 (10)	2.786	2.960 (5)	25	P L	
HOU	0.243	1.067	0.789 (5)	1.313	1.608 (15)	1.804	20	I E S	
	10	10	5	10	15	10			
	D E M A N D S								

TABLE 1: FULL TRANSPORTATION PROBLEM

(Optimal solution shown in parentheses; optimal value = \$116,005)

Grouping Pittsburgh, New York and Boston together is proposed as the first trial aggregation. This leads to the reduced problem shown in Table 2 (again, disregard the optimal solution shown). Notice that the demand of PIT/NY/BOS is just the sum of the individual demands, and that the unit

costs in this column are weighted pro rata demand. For instance, the SEA → PIT/NY/BOS unit cost is:

$$\frac{10}{35} \times 2.465 + \frac{15}{35} \times 2.815 + \frac{10}{35} \times 2.976 = 2.761 \text{ \$/CWT.}$$

(It is an elementary exercise to show that this pro rata demand aggregation is equivalent to the one studied in Sec. 1.)

	DAL	CHI	ATL	PIT/NY/BOS		
SEA	2.078	2.013 (10)	2.618	2.761 (5)	15	S U P P L I E S
LA	1.387 (10)	2.054	2.182	2.733 (15)	25	
HOU	.243	1.067	.789 (5)	1.580 (15)	20	
	10	10	5	35		

D E M A N D S

TABLE 2: FIRST AGGREGATION

(Optimal solution shown in parentheses; optimal value = \$116,445)

The a priori error bound associated with this aggregation is

$\epsilon_{\text{PIT,NY,BOS}} = \500 . From past experience, management expects the total transportation bill to be in the general vicinity of \$100,000. This provides a basis for deciding whether the proposed aggregation is satisfactory.

Suppose that \$500, an error of about 1/2 of 1%, is judged excessively high -- and hence that a less coarse aggregation such as New York/Boston seems advisable. This leads to the problem shown in Table 3.

	DAL	CHI	ATL	PIT	NY/BOS	
SEA	2.078	2.013 (10)	2.618	2.465	2.879 (5)	15
LA	1.387 (10)	2.054	2.182	2.426 (10)	2.856 (5)	25
HOU	.243	1.067	.789 (5)	1.313	1.686 (15)	20
	10	10	5	10	25	

D E M A N D S

TABLE 3: SECOND AGGREGATION

(Optimal solution shown in parentheses, optimal value = \$116,170)

The a priori error bound for this second aggregation is smaller:

$\epsilon_{NY,BOS} = \$200$. To show the calculation in detail, we first express the general formula (9) in terms of unit transportation costs c_{kl} , i.e., we write d_{kl} as $q_l c_{kl}$ to obtain an equivalent representation

$$(9A) \quad \epsilon_l = \sum_{l' \in L} q_{l'} \max_k \left\{ \frac{\sum_{l' \in L} q_{l'} c_{kl'}}{\sum_{l' \in L} q_{l'}} - c_{kl} \right\}$$

Now putting $L = \{NY, BOS\}$, we have

$$\begin{aligned} \epsilon_{NY,BOS} &= q_{NY} \max_{k=1,2,3} \left\{ \frac{q_{NY} c_{k,NY} + q_{BOS} c_{k,BOS}}{q_{NY} + q_{BOS}} - c_{k,NY} \right\} \\ &+ q_{BOS} \max_{k=1,2,3} \left\{ \frac{q_{NY} c_{k,NY} + q_{BOS} c_{k,BOS}}{q_{NY} + q_{BOS}} - c_{k,BOS} \right\} \\ &= 15,000 \max \{ 2.879 - 2.815, 2.856 - 2.786, 1.686 - 1.608 \} \\ &+ 10,000 \max \{ 2.879 - 2.976, 2.856 - 2.960, 1.686 - 1.804 \} \\ &= (15,000) (0.078) + (10,000) (-0.097) = \$200. \end{aligned}$$

Assuming that this smaller error bound is deemed acceptable, the second aggregation would be accepted as sufficiently accurate.

This concludes the miniature illustration of the trial-and-error process by which different aggregations can be proposed and evaluated until one is found with an acceptable compromise between parsimony and exposure to modeling error.

It may also be of interest to use this illustration to exemplify the assertions of the Customer Aggregation Theorem in somewhat greater detail. The optimal solution of the full transportation problem and its two aggregations are shown in Tables 1-3. Evidently the inequalities of the theorem turn out to be satisfied:

$$\underbrace{\text{optimal value of full problem}}_{116,005} \leq \underbrace{\text{optimal value of PIT/NY/BOS aggregation}}_{116,445} \leq \underbrace{\text{optimal value of full problem}}_{116,005} + \underbrace{\epsilon_{\text{PIT/NY/BOS}}}_{500} = 116,505$$

$$\underbrace{\text{optimal value of full problem}}_{116,005} \leq \underbrace{\text{optimal value of NY/BOS aggregation}}_{116,170} \leq \underbrace{\text{optimal value of full problem}}_{116,005} + \underbrace{\epsilon_{\text{NY/BOS}}}_{200} = 116,205$$

Moreover, by virtue of the nature of pro rata demand aggregation, the solution to each aggregated problem can easily be disaggregated into a feasible solution to the full problem with the same cost. For instance, the SEA - NY/BOS flow of 5,000 CWT in the second aggregation would be disaggregated into

$$\frac{15}{25} \times 5000 = 3000 \text{ CWT to New York}$$

$$\frac{10}{25} \times 5000 = 2000 \text{ CWT to Boston}$$

without changing the associated transportation cost ($5000 \times 2.879 = 3000 \times 2.815 + 2000 \times 2.976$). The other flows to NY/BOS would be disaggregated similarly. The resulting disaggregated feasible solution to the full problem is suboptimal to within its respective a priori error bound. For the second aggregation, for instance, the suboptimality of $\$116,170 - 116,005 = \165 is under the a priori bound of $\$200$.

3. Sufficient Condition for Zero Aggregation Error

It is of interest to examine the conditions under which (9) yields $\epsilon_L = 0$. One easy sufficient condition is as follows.

Corollary. Suppose that the d_{kl} 's can be written in the factored form

$$(10) \quad d_{kl} = q_l \lambda_k \quad \text{for all } kl \text{ with } l \in L$$

for some suitable set of λ_k 's. Then $\epsilon_L = 0$. (However, (10) is not a necessary condition for $\epsilon_L = 0$.)

This agrees with one's expectation that ϵ_L should be 0 when the customers of L are all situated close together geographically and are "similar" in terms of demand type, for then λ_k would have a natural interpretation in terms of the cost per unit quantity of satisfying the needs of any customer in L from facility k . More specifically, suppose that d_{kl} is composed of an "acquisition" cost α_k $\$/unit$ plus a transportation cost based on a rate β $\$/unit-mile$. If the distance from facility k to any customer in L is virtually the same, say d_k miles, then

$$(11) \quad d_{kl} = q_l (\alpha_k + \beta d_k) \quad \text{for all } kl \text{ with } l \in L$$

and (10) holds with $\lambda_k = \alpha_k + \beta d_k$.

4. Extension to Several Aggregation Subsets

A natural extension of the Customer Aggregation Theorem addresses the case where several subsets of customers are to be aggregated simultaneously. Let L_1, \dots, L_H be disjoint subsets of customer indices such that each subset L_h individually satisfies (7). The analog of (6) is

(6)' for $h=1, \dots, H$ and all k , the y_{kl} 's are identical over $l \in L_h$.

It is not difficult to show that the analog of (8) still holds, namely

(8)' $v[\text{problem (1)-(5)}] \leq v[\text{problem (1)-(5), (6)'}] \leq v[\text{problem (1)-(5)}] + \epsilon'_L$,

where

$$(9)' \quad \epsilon'_L = \Delta \sum_{h=1}^H \sum_{l \in L_h} \text{Max}_k \left\{ \frac{q_l \sum_{l' \in L_h} d_{kl'}}{\sum_{l' \in L_h} q_{l'}} - d_{kl} \right\}.$$

This is exactly the same result as would be obtained from H sequential applications of the Customer Aggregation Theorem in its original form.

5. Extension to a Two-Stage Model

Another natural extension is to the case of a two-stage distribution system. The natural two-stage version of problem (1)-(5) is:

(12) Minimize $\sum_{jk} c_{jk} x_{jk} + \sum_{kl} d_{kl} y_{kl} + \sum_k F_k(\sum_l q_l y_{kl}, z_k)$
 x, y, z

(13) subj. to $\underline{s}_j \leq \sum_k x_{jk} \leq \bar{s}_j$, all j

(14) $\sum_j x_{jk} = \sum_l q_l y_{kl}$, all k

(2) $\sum_k y_{kl} = 1$, all l

(3) $\underline{v}_k z_k \leq \sum_l q_l y_{kl} \leq \bar{v}_k z_k$, all k

(15) $x_{jk} \geq 0$, all jk

(4), (5) $0 \leq y_{kl} \leq 1$, all kl ; $z_k = 0, 1$ for all k ; $z \in Z$.

The following new interpretations are appropriate:

- j indexes the sources which supply the facilities
- x_{jk} a variable giving the annual amount of supplies obtained by facility k from source j
- c_{jk} unit cost of procurement or production plus transportation associated with the flow x_{jk}
- \underline{s}_j (\bar{s}_j) a lower (upper) limit on the annual amount of supplies procured from source j .

The interpretation of the two-stage problem should be evident. Constraints (14) amount to an annual material balance requirement at each facility.

It is easy to show that the Customer Aggregation Theorem holds without change. The only alteration needed in the proof given in the Appendix involves the addition of $\bar{x}_{jk} = x_{jk}^0$, all jk , to (A4).

6. Designing Aggregations by Clustering

Up to this point we have taken the viewpoint that a priori bounds are useful for evaluating the comparative merits of alternative customer aggregations on a case by case basis. An obvious next step would be to attempt to automate the aggregation design process by seeking the coarsest possible customer aggregation with an a priori error bound no larger than some pre-specified limit.

This leads to a well-defined but exceedingly difficult combinatorial optimization problem. Fortunately, truly optimal solutions are quite unnecessary and so it is appropriate to seek good heuristic techniques.

The methods of cluster analysis [1] are attractive as a source of heuristic techniques for customer aggregation. A hierarchial (that is, parametric) clustering approach starting with as many clusters as original customers and then combining clusters one at a time, perhaps with periodic individual reclassification of customers, appears to hold promise. The result would be an approximate tradeoff curve between the number of customer clusters and the magnitude of the à priori bound given by (9)'.

The chosen measure of association or distance between two compatible customer clusters, say L_i and L_j , plays a vital role. Perhaps the most natural measure is this one:

$$(16) \quad d(L_i, L_j) \triangleq \epsilon_{L_i \cup L_j} - (\epsilon_{L_i} + \epsilon_{L_j}).$$

The additivity property of ϵ_L developed in Sec. 4 is essential in justifying this measure in that it causes the influence of all other clusters to cancel out.

An attempt is currently under way to develop and test clustering techniques based on (16) as a means of at least partially automating the customer aggregation aspect of distribution model design.

References

1. Anderberg, M.R., Cluster Analysis for Applications, Academic Press, 1973.
2. Geoffrion, A.M., "A Guide to Computer-Assisted Methods for Distribution Systems Planning," Sloan Management Review, Winter 1975, 17-41.
3. Geoffrion, A.M., "A Priori Error Bounds for Procurement Commodity Aggregation in Logistics Planning Models," Naval Research Logistics Quarterly, forthcoming 1977.
4. Morris, W.T., "On the Art of Modeling," Management Science, 13, 12 (August 1967), B-707 to B-717.

APPENDIX:

PROOF OF THE CUSTOMER AGGREGATION THEOREM

The proof is based upon the following fundamental result.

Lemma 1 (Restrictive Approximation). Consider a general mathematical programming problem

$$(P) \quad \underset{x}{\text{Minimize}} \quad f(x) \quad \text{subject to } x \in X$$

and also the following restrictive approximation to it

$$(Q) \quad \underset{x}{\text{Minimize}} \quad f(x) \quad \text{subject to } x \in X \cap \bar{X},$$

where X and \bar{X} are both subsets of the same set. Assume that X is not empty and that a *Feasibility Recovery Rule*

is known which associates to every point x^0 in X some point $\bar{x}(x^0)$ in $X \cap \bar{X}$ in such a manner that, for some scalar ϵ_1 ,

$$(A1) \quad f(\bar{x}(x^0)) \leq f(x^0) + \epsilon_1 \quad \text{for all } x^0 \in X.$$

Then

$$(A2) \quad v[P] \leq v[Q] \leq v[P] + \epsilon_1$$

and every ϵ_2 -optimal solution x^0 of (Q) is $(\epsilon_1 + \epsilon_2)$ -optimal in (P),

where $v[\cdot]$ denotes the infimal value of problem $[\cdot]$.

Proof. We have

$$v[Q] \leq \underset{x \in X}{\text{Infimum}} \quad f(\bar{x}(x^0)) \leq v[P] + \epsilon_1,$$

where the first inequality follows from the fact that $\bar{x}(x^0) \in X \cap \bar{X}$ and the second inequality follows from (A1). This, with the evident fact that $v[P] \leq v[Q]$, proves (A2). The other desired conclusion requires demonstrating

$$(A3) \quad f(x^0) \leq v[P] + \epsilon_1 + \epsilon_2,$$

since x^0 clearly is feasible in (P). This follows directly from (A2) and the hypothesis

$$f(x^0) \leq v[Q] + \epsilon_2.$$

The most fruitful applications of this result are those in which a feasibility recovery rule can be devised with a small associated ϵ_1 .

In the context of problem (1)-(5) and its aggregated version (1)-(6), the following rule appears to be the most natural one.

Lemma 2. Let (y^0, z^0) be any feasible solution of problem (1)-(5), and let L be any subset of demand points satisfying (7). Make the definitions

$$(A4) \quad \begin{cases} \bar{y}_{kl} \triangleq y_{kl}^0 & , \quad \text{all } kl \text{ with } l \notin L \\ \bar{y}_{kl} \triangleq \epsilon_k & , \quad \text{all } kl \text{ with } l \in L \\ \bar{z}_k \triangleq z_k^0 & , \quad \text{all } k \end{cases}$$

where, for each k such that kl exists for $l \in L$, we put

$$(A5) \quad \epsilon_k \triangleq \frac{\sum_{l \in L} q_l y_{kl}^0}{\sum_{l \in L} q_l}.$$

Then (\bar{y}, \bar{z}) as so defined is feasible in the aggregated problem (1)-(6) and

$$(A6) \quad \left\{ (1) \text{ evaluated at } (\bar{y}, \bar{z}) \right\} - \left\{ (1) \text{ evaluated at } (y^0, z^0) \right\} \\ = \sum_{\substack{kl \\ l \in L}} \left\{ \frac{q_l \sum_{l' \in L} d_{kl'}}{\sum_{l' \in L} q_{l'}} - d_{kl} \right\} y_{kl}^0$$

Proof. First we verify the feasibility of (\bar{y}, \bar{z}) in problem (1)-(6).

Constraint (2) holds for $l \notin L$ because the y_{kl} 's involved were not changed. For $l \in L$, we have

$$\begin{aligned} \sum_k \bar{y}_{kl} &= \sum_k \xi_k = \sum_k \sum_{l' \in L} q_{l'} y_{kl}^0 / \sum_{l' \in L} q_{l'} \\ &= \sum_{l' \in L} q_{l'} (\sum_k y_{kl}^0) / \sum_{l' \in L} q_{l'} = 1, \end{aligned}$$

where the last equality follows from

$$\sum_k y_{kl}^0 = 1 \text{ for all } l'.$$

This verifies that (2) holds for $l \in L$ as well as for $l \notin L$. To verify

(3) we need the result

$$(A7) \quad \sum_l q_l \bar{y}_{kl} = \sum_l q_l y_{kl}^0 \quad \text{for all } k,$$

which can be seen as follows:

$$\begin{aligned} \sum_l q_l \bar{y}_{kl} &= \sum_{l \notin L} q_l \bar{y}_{kl} + \sum_{l \in L} q_l \bar{y}_{kl} \\ &= \sum_{l \notin L} q_l y_{kl}^0 + \frac{\sum_{l \in L} q_l \sum_{l' \in L} q_{l'} y_{kl}^0}{\sum_{l' \in L} q_{l'}} \\ &= \sum_{l \notin L} q_l y_{kl}^0 + \sum_{l' \in L} q_{l'} y_{kl}^0. \end{aligned}$$

We remark that the need for (A7) to hold furnished the primary motivation for the choice (A5). It follows easily from (A7) that (3) holds at

(\bar{y}, \bar{z}) . Constraint (4) holds for \bar{y} because (4) holds for y^0 and each ξ_k is just a convex combination of certain y_{kl}^0 's. Constraint (5) holds because $\bar{z} = z^0$. The very nature of (A4) implies that (6) holds at \bar{y} .

This establishes that (\bar{y}, \bar{z}) is feasible in the aggregated problem (1)-(6).

Now subtract (1) evaluated at (y^0, z^0) from (1) evaluated at (\bar{y}, \bar{z}) .

The result is

$$\sum_{kl} q_{kl} (\bar{y}_{kl} - y_{kl}^0) + \sum_k F_k (\sum_l q_l \bar{y}_{kl} - \bar{z}_k) - \sum_k F_k (\sum_l q_l y_{kl}^0 - z_k^0)$$

$$\begin{aligned}
 &= \sum_{k \in L} d_{kl} (\xi_k - y_{kl}^{\circ}) \\
 &= \sum_{k \in L} d_{kl} \left(\frac{\sum_{l' \in L} q_{l'} y_{kl'}^{\circ}}{\sum_{l' \in L} q_{l'}} - y_{kl}^{\circ} \right) \\
 &= \sum_{k \in L} \left(\frac{q_k \sum_{l' \in L} d_{kl'}}{\sum_{l' \in L} q_{l'}} - d_{kl} \right) y_{kl}^{\circ} ,
 \end{aligned}$$

where the first equality follows from (A7), the second from (A5), and the third from a rearrangement of terms. This demonstrates (A6).

The Customer Aggregation Theorem follows easily from Lemmas 1 and 2 upon making the obvious identifications. The one additional fact needed is that the difference expression (A6) is bounded above by ϵ_L as defined in (9), since y° must satisfy (2) and (4).