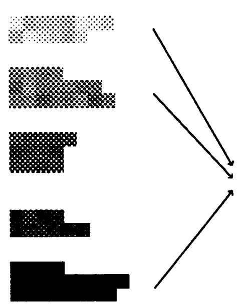
Report on CAT Experiments

By

Mark H. Hansen and Kenneth W. Wachter

Technical Report No. 362
July 1992

Department of Statistics
University of California
Berkeley, California 94720



Report on CAT Experiments

by Mark H. Hansen and Kenneth W. Wachter

April 29, 1992

I. Introduction

The CAT experiments were introduced by Kenneth Wachter in his report to the Secretary of Commerce Robert Mosbacher in June of 1991 as a means for assessing the impact which failures of the homogeneity assumption might have in addition to other sources of systematic error when adjustments are carried down to local levels. The homogeneity assumption is the assumption that all people in the same Post-Enumeration Survey (PES) post-strata have the same probabilities of Census omission and erroneous enumeration, whatever their location. Errors due to the failure of this assumption are additional errors beyond those taken into account in the Bureau's Total Error Model or in the Bureau's error margins for adjusted sizes of places based on the PES.

The CAT experiments address the question of what levels of error would result in the absence of systematic errors in the PES when adjustments are carried down to local areas like Census blocks or districts. In these experiments, local areas are represented by blocks in the PES sample within the area, not by all Census blocks. The CAT experiments are computer simulation experiments. The computer creates a replica for the outcomes of each stage of the adjustment process: generating a Census, conducting a Post-Enumeration Survey, and calculating the adjustment factors. In the context of the computer experiment, unlike the real world, we know the true underlying population and can evaluate the relative accuracy of the Census and adjusted counts. The acronym CAT stands for "Census-Adjustment-Truth," and it is from our ability to compare Census counts and Adjusted counts against the True population that these simulations derive their name.

The data on which we will base our simulations was taken directly from the 1990 Post-Enumeration Survey. The smallest unit of aggregation for our PES data is referred to as a block cluster. A block cluster is a contiguous area containing a few hundred people. In cities, a block cluster is usually a city block or half-block or a few contiguous blocks. In the country, it is an area of comparable population. From the Advisory Use File of individual-level PES data, with fields deleted to protect confidentiality, we tabulated block cluster-level PES results from block clusters in each of ten different geographic areas. The resulting sets of block clusters are referred to as sites. Further, each of our geographic areas

Site 1	Middle Atlantic Division, Central City, New York City (from Post-Stratum Groups 4 to 9) 23 District Offices, 125 block clusters
Site 2	East North Central Division, Central City, Chicago (from Post-Stratum Groups 64 to 67) 11 District Offices, 112 block clusters
Site 3	East North Central Division, Central City, Detroit (from Post-Stratum Groups 64 to 67) 6 District Offices, 70 block clusters
Site 4	West South Central Division (including Texas), Central and Non-Central Cities in Type II Metropolitan Areas (from Post-Stratum Groups 56, part of 57, part of 58, 59) 23 District Offices, 104 block clusters
Site 7	Pacific Division, Non-Central Cities in Metropolitan Areas other than Los Angeles (from Post-Stratum Groups 104 to 108) 52 District Offices, 135 block clusters

Table 1

contains a number of Census districts. Those block clusters in our sample that fall into the same Census district are combined to form district chunks. In what follows, when we refer to "blocks" or to "block-level data," we actually mean block clusters and block-cluster-level data. Similarly, when we refer to "districts" we are actually referring to district chunks.

Of our ten sets of PES blocks clusters, or sites, we will concentrate on five. Their complete descriptions are given above. Because our sites 2 and 3 are actually combined in the PES, we will create a new site from the cities of Detroit and Chicago, which we will refer to as Detago. For convenience, we will also refer to site 1 as New York City, site 4 as Arlington, and to site 7 as Berkeley.

As mentioned previously, our PES data have been cross-classified by block cluster and demographic categories. In what follows, we will consider 48 such demographic categories, or demoids, based on race, sex and age. Unlike the previous CAT simulations, we consider four different race groups and 6 different age groups. The various categories are given in Tables 2, 3 and 4. Each combination of the race/sex/age categories listed in these tables is referred to as a demoid. The demoids are labeled by the integers from 1 to 48, using the rule

$$\text{Demoid} = 12*(\text{Race} -1) + 6*(\text{Sex}-1) + \text{Age},$$

where Race, Sex, and Age are the numeric labels given in Tables 2, 3 and 4.

	Race		Sex		Age
1	White	1	Male	1	Ages 0-9
2	Black	2	Female	2	Ages 10-19
3	Hispanic			3	Ages 20-29
4	Asian			4	Ages 30-44
				5	Ages 45-64
				6	Ages 65+

Tables 2, 3 and 4

A word is in order about the Detago Site. It is important to bear in mind that despite the differences between Detroit and Chicago, the PES treats these cities as a single unit. People in a given age, sex, race and tenure group share the same adjustment factor, regardless of the city in which they reside. Detroit and Chicago have been combined for these CAT simulations in order to investigate the effects that such combining may have in the actual PES. Note that the demoids do not exactly correspond to post-strata. Hispanics from other cities are grouped with those from Detroit and Chicago in the PES post-strata. Renters and owners are combined in the demoids used for the simulations in this and in the other sites.

The rest of this report introduces both the design and the rationale behind the current implementation of the CAT simulations. Replicating the adjustment process begins with simulating the results from a Census. Starting from a hypothetical true count of people, this is equivalent to randomly deciding who will be correctly enumerated, who will be a gross omission, and who will be erroneously enumerated. This process is introduced in Section II, along with a complete discussion of the techniques we have explored for estimating the probabilities associated with each of these three events. In Section III, we introduce the sampling scheme by which we will create the hypothetical true population counts used to simulate our Census results. The details from Section II and III are combined in Section IV, when we present a complete description of the current implementation of our CAT simulations. The last Section is reserved for presenting our results. We have run complete district- and block-level simulations for each of New York, Arlington, Berkeley and Detago.

II. Modeling of Undercount Rates

The first step of our CAT simulations involves generating a random Census from a table of hypothetical true population counts that have been cross classified by district and demoid. Our main concern at this stage is that each randomly generated Census exhibit patterns of gross omissions and erroneous enumerations across both districts and demographic groups that are similar to those that we observe in our PES data.

Given a table of hypothetical true counts, we generate a Census by randomly labeling each person in the table as either being missed, correctly enumerated, or erroneously enumerated by the Census. Clearly, our ability to reproduce patterns of gross omissions and erroneous enumerations in such random realizations depends on the probabilities we assign to each of these three outcomes. In this Section, we present four different techniques, denoted FIT1 through FIT4, for estimating these probabilities from our PES data. While each incorporates demographic information, only the first assumes homogeneity across districts, which results in estimates that depend solely on a person's demographic group. The remaining three estimation techniques differ from each other in the way they treat both district and demographic information. Extensive simulation results based on FIT1 and FIT3 will be presented in Section V. Each of FIT1 through FIT4 have been applied to both district- and block cluster-level data. However, results from FIT3 are broadly typical of results from

FIT2, FIT3 and FIT4. The differences among the three are interesting, but are too small to be a major concern in this report.

Let go(b,d) represent the chance that a person in district b and demographic group d is missed by the Census and let ee(b,d) represent the chance that the same person is erroneously enumerated. Therefore, the probability that this person is correctly enumerated by is given by 1-go(b,d)-ee(b,d). Without ignoring the fact that erroneous enumerations occur in the Census for a variety of different reasons, our simulations will treat erroneous enumerations in a way that is easiest to interpret if they are viewed as being duplicates. Thus, at the most basic level, we are interested in calculating the chance associated with a person being counted once, twice or not at all by the Census. These three possibilities are presented in Table 5.

	Census 0	Census 1	Census 2
Probability	go(b,d)	1-go(b,d)-ee(b,d)	ee(b,d)

Table 5

Each of the four fitting techniques begins with the calculation of some form of demoid "effect." By demoid effects, we mean estimates $\hat{go}(d)$ and $\hat{ee}(d)$ of go(b,d) and ee(b,d), respectively, that are functions of demoid d alone. Let C(b,d), E(b,d), P(b,d), and M(b,d) denote the Census count, the number of erroneous enumerations, the P-sample count and the number of P-sample matches tallied for a given district b and demoid d. Further, let C(b), E(b), P(b) and M(b) denote similar quantities pooled across demoids for a given district b, and let C(d), E(d), P(d) and M(d) denote the various counts summed across districts for a fixed for a fixed demoid d. We then calculate our demoid effects as

$$\hat{go}(d) = \frac{P(d) - M(d)}{P(d)}, \qquad \text{and} \qquad \hat{ee}(d) = \frac{E(d)}{C(d) - E(d)} \left(\frac{M(d)}{P(d)}\right),$$

respectively. In the second expression, we are dividing the number of erroneous enumerations by the Bureau's Dual System Estimate of the true count for the demoid.

FIT1 takes $\hat{go}(d)$ and $\hat{ee}(d)$ as estimates of go(b,d) and ee(b,d). The assumption that the probability of being a gross omission or an erroneous enumeration depends only on a person's demographic group is exactly the Census Bureau's homogeneity assumption upon which their adjustment procedure is based.

Within each demoid in a given site, the gross omissions and erroneous enumerations tend to be concentrated in a subset of districts rather than spread evenly throughout the site. To quantify this behavior, we introduce gross omission and erroneous enumeration rates. The gross omission rate, or go-rate, for a district b and a demoid d is defined as (P(b,d) - M(b,d))/P(b,d), the number of gross omissions divided by the P-sample count of people in district b and demoid d. Similarly, the go-rate for a district b is given by (P(b) - M(b))/P(b), and the go-rate for a demoid d is (P(d) - M(d))/P(d). The erroneous enumeration rate, ee-rate, for a district b and a demoid d is defined as E(b,d)/C(b,d), the number of erroneous enumerations divided by the Census count of people in district b and demoid d. As one might expect, the ee-rate for a district b is given by E(b)/C(b), and the ee-rate for a demoid d is E(d)/C(d). The go- and ee-rates for other classifications of people in a site are defined similarly.

As an example, in Figure 1, we present tables of go- and ee-rates calculated by district and race group for the combined PES site Detago. In order to downplay the effects of sampling error in small cells, entries in these tables that are computed from Census or P-sample counts of fewer than 25 people are left blank. The first six districts in these tables are located in Detroit, while the last eleven belong to Chicago. Differences in the pattern of go-rates in the two cities are easily identified. For example, the go-rate among blacks living in Detroit is relatively constant across districts, while in Chicago the go-rate for blacks varies considerably from district to district. Similarly, we observe differences in the pattern of ee-rates between these two cities. Recall that we have combined Detroit and Chicago into one Detago site not because of any similarity between them, but because they are combined in the PES.

	White	Black	Hisp	Asian
1	0.10	0.13		0.16
2	0.15	0.12		
3		0.11		
4	0.06	0.11		
5		0.09		
6	0.06	0.11	0.12	
7	0.23	0.24		
8	0.16	0.26	0.21	
9	0.04	0.33	0.13	
10	0.03	0.16	0.14	
11	0.08		0.24	
12	0.06	0.14	0.27	
13	0.21	0.21	0.23	0.07
14	0.09	0.16	0.17	
15	0.04		0.10	0.00
16		0.20		
17	0.04	0.21		

Go-rate in Detago by District and Race

	White	Black	Hisp	Asian
1	0.10	0.11		0.18
2		0.18		
3		0.10		
4	0.05	0.10		
5		0.07		
6	0.07	0.06	0.00	
7	0.03	0.17		
8		0.15	0.13	0.23
9	0.10	0.17	0.09	0.15
10	0.04	0.08	0.06	
11	0.05		0.10	
12	0.03	0.07	0.16	
13	0.12	0.12	0.15	0.18
14	0.05	0.07	0.19	0.19
15	0.02		0.04	0.03
16		0.08		
17	0.03	0.29		

Ee-rate in Detago by District and Race

Figure 1: Tables of go- and ee-rates by District and Race Group in Detago.

One way to include observed "heterogeneity" between districts in our estimates of go(b,d) and ee(b,d) is to scale our demoid effects $\hat{go}(d)$ and $\hat{ee}(d)$ from FIT1 by district-specific factors. In so doing, we would be assuming that while we have observed differences in the patterns of go- or ee-rates between districts, their relative patterns among demoids in a given district is the same for all districts. This assumption leads us to our second estimation technique, FIT2. To calculate our scale factors, we begin by using $\hat{go}(d)$ and $\hat{ee}(d)$ to predict the expected number of gross omissions and erroneous enumerations, respectively, for a given district. These predictions are are defined as

$$\sum_d P(b,d)\hat{go}(d), \quad \text{and} \quad \frac{P(b)}{M(b)} \sum_d (C(b,d)-E(b,d))\hat{ee}(d),$$

where each of the sums is over all 48 demoids. Observe that in the second expression above we are again using the Dual System Estimate, this time to estimate the true population in a district. We then calculate our district-specific scale factors by comparing these predictions to the observed number of gross omissions and erroneous enumerations for each district. Let $\hat{go}(b)$ and $\hat{ee}(b)$ denote these scale factors and set

$$\hat{go}(b) = \frac{P(b)-M(b)}{\sum_d P(b,d)\hat{go}(d)}, \quad \text{and} \quad \hat{ee}(b) = \frac{M(b)}{P(b)} \frac{E(b)}{\sum_d (C(b,d)-E(b,d))\hat{ee}(d)}.$$

Finally, we arrive at our estimates of the probabilities of gross omission and erroneous enumeration for a demoid d in a district b by setting

$$\hat{go}(b,d) = \hat{go}(b)\hat{go}\,(d) \quad \text{and} \quad \hat{ee}(b,d) = \hat{ee}(b)\hat{ee}\,(d).$$

Predictions of the number of gross omissions and erroneous enumerations for a given district based on $\hat{go}(b,d)$ and $\hat{ee}(b,d)$ now exactly match the observed number of gross omissions and erroneous enumerations in the district.

Our third and fourth fitting techniques also incorporate district information through scale factors similar to those calculated in FIT2. However, we now introduce a district and demoid "interaction." That is, we will now allow for different relative patterns of go- and/or ee-rates among demoids from different districts. This will be done by using go- and ee-rates to classify the districts in a given site as being either "light" (contributing a relatively small number of gross omissions or erroneous enumerations) or "heavy" (contributing a relatively large number of gross omissions or erroneous enumerations), and fitting different demoid effects based on this classification.

Define the "action" in a given district as the sum of its go-rate and ee-rate. For each site, we begin by sorting the districts so that those with the largest action values come first. Taking the districts in this order, we classify as "heavy" those districts that account for half of the total erroneous enumerations and gross omissions in the site. The other districts are labeled as "light."

Let the subscript h on any of the quantities $C(d)$, $E(d)$, $P(d)$ and $M(d)$ indicate that rather than pooling these various counts across all districts for a fixed demoid d, we have instead summed only across districts that have been labeled as heavy. Similarly, the subscript l ("ell") indicates that these counts have been summed across only the light districts. As noted above, demoid effects for the probability of gross omission or erroneous enumeration are estimated separately for heavy and light districts in FIT3 and FIT4. In each case, the demographic effects from from heavy districts are found using

$$\hat{go}_h(d) = \frac{P_h(d) - M_h(d)}{P_h(d)}, \qquad \text{and} \qquad \hat{ee}_h(d) = \frac{E_h(d)}{C_h(d) - E_h(d)} \left(\frac{M_h(d)}{P_h(d)}\right).$$

These are exactly the same expressions we derived for the demoid effects in FIT1, except that now the various counts are summed across only heavy districts.

FIT3 and FIT4 differ only in their treatment of the demographic effects in light districts. In FIT3, we use expressions similar to the ones above to derive $\hat{go}_l(d)$ and $\hat{ee}_l(d)$. That is,

$$\hat{go}_l(d) = \frac{P_l(d) - M_l(d)}{P_l(d)}, \qquad \text{and} \qquad \hat{ee}_l(d) = \frac{E_l(d)}{C_l(d) - E_l(d)} \left(\frac{M_l(d)}{P_l(d)}\right).$$

In FIT4, however, we assume that the demoid effects from the light districts are constant: for each demoid d, $\hat{go}_l(d) = \hat{go}_l$ and $\hat{ee}_l(d) = \hat{ee}_l$ where

$$\hat{go}_l = \frac{\sum_d (P_l(d) - M_l(d))}{\sum_d P_l(d)}, \qquad \text{and} \qquad \hat{ee}_l = \frac{\sum_d E_h(d)}{\sum_d (C_h(d) - E_h(d))} \left(\frac{\sum_d M_h(d)}{\sum_d P_h(d)}\right).$$

In this case, the demographic effects from light districts can be viewed as background rates that reflect the average probabilities of gross omission and erroneous enumeration across demoids in the light districts.

Regardless of which of these two estimates we use for the demoid effects in the light districts, we complete each of our last two estimation schemes, FIT3 and FIT4, by again determining district-specific scale factors from predictions of the number of gross omissions and erroneous enumerations made using only our demoid effects. For example, if we assume that a district b is heavy, we calculate

$$\hat{go}(b) = \frac{P(b)\text{-}M(b)}{\sum_d P(b,d)\hat{go}_h(d)}, \quad \text{and} \quad \hat{ee}(b) = \frac{M(b)}{P(b)} \frac{E(b)}{\sum_d (C(b,d)\text{-}E(b,d))\hat{ee}_h(d)},$$

and arrive at estimates of the probabilities of gross omission and erroneous enumeration for a demoid d in heavy district b by setting $\hat{go}(b,d) = \hat{go}(b)\hat{go}_h(d)$ and $\hat{ee}(b,d) = \hat{ee}(b)\hat{ee}_h(d)$. We replace $\hat{go}_h$ and $\hat{ee}_h$ by $\hat{go}_l$ and $\hat{ee}_l$, respectively, in these expressions to obtain $\hat{ee}(b,d)$ and $\hat{go}(b,d)$ for light districts. To summarize, FIT3 and FIT4 estimate different demoid effects for heavy and light districts. This is done to allow for different relative patterns of go- and/or ee-rates among demoids from heavy and light districts. Then, as was done in FIT2, we calculate district-specific scale factors so that predictions of the number of gross omissions and erroneous enumerations in a given district based on the resulting estimated probabilities exactly match their observed numbers in the given district.

When our simulations are run at the level of blocks rather than districts, we extend of FIT1 through FIT4 to estimate go(b,d) and ee(b,d), where these now represent the chance that a person living in block b with a set of demographic characteristics d is either a gross omission or an erroneous enumeration. Because they depend only on demographic information, our estimates from FIT1 will not change when we move from district-level to block-level simulations.

While we have run our CAT simulations using probabilities estimated from each of our four fitting techniques, in what follows, we will concentrate mainly on results obtained from FIT1 and FIT3. We will now consider differences between these two fits with respect to their ability to capture patterns of gross omissions and erroneous enumerations in a given site. We begin by examining results from the combined PES site Detago. To this site we fit our two different models for go(b,d) and ee(b,d). Next, we use these probabilities to predict the number of gross omissions and erroneous enumerations in each of four race groups (White, Black, Hispanic and Asian) within the 17 districts in the site. The residuals (defined as actual minus predicted) from these predictions are displayed in the following perspective plots as heights above a series of 17 by 4 arrays. Figure 2 presents the results for gross omissions. Here, over the table on the right we plot the the residuals from FIT3, which includes district effects and the district by demoid interaction, while over the table on the left we plot the residuals from FIT1, which recognizes only demographic effects. In these tables, the 17 districts are ordered according to the action in each district (the heavy districts appearing at the top of each table and the light districts appearing at the bottom). Taken from left to right, the four columns of each table represent the White, Black, Hispanic and Asian race groups, respectively. It is clear from these tables that predictions made by considering only demoid effects tend both to underestimate the number gross omissions in heavy districts, and to overestimate their number in light districts. Including district effects as we did in our third fit reduces this pattern considerably. In Figure 3, we present a similar perspective plot of the residuals from our two sets of predictions for the number of erroneous enumerations in each district by race group. Once again, we observe a pattern of underestimating then overestimating the number of erroneous enumerations for a fixed race group as we move from heavy to light districts in the table on the right. More conventional pairwise plots of predicted versus actual gross omissions and erroneous enumerations across districts for a fixed race group are included in the appendix. In each of these plots, heavy districts are marked with a "1," light districts are marked with a "2," and the dashed line has unit slope and zero intercept.

The impact of district effects on our fits in Detago can be explained by the dramatically different patterns of gross omissions and erroneous enumerations across race groups that we observed previously. Similar results are found when we consider site 1 (New York). In Figures 4 and 5, we present perspective plots of the residuals from predictions made from FIT1 and FIT3 for the number of gross omissions and erroneous enumerations by district and race group taken from New York. Again, the 23 districts in this site are arranged in the tables in order of decreasing action, and the columns represent the White, Black, Hispanic and Asian race groups. Simple pairwise plots for this site are also given in the appendix.

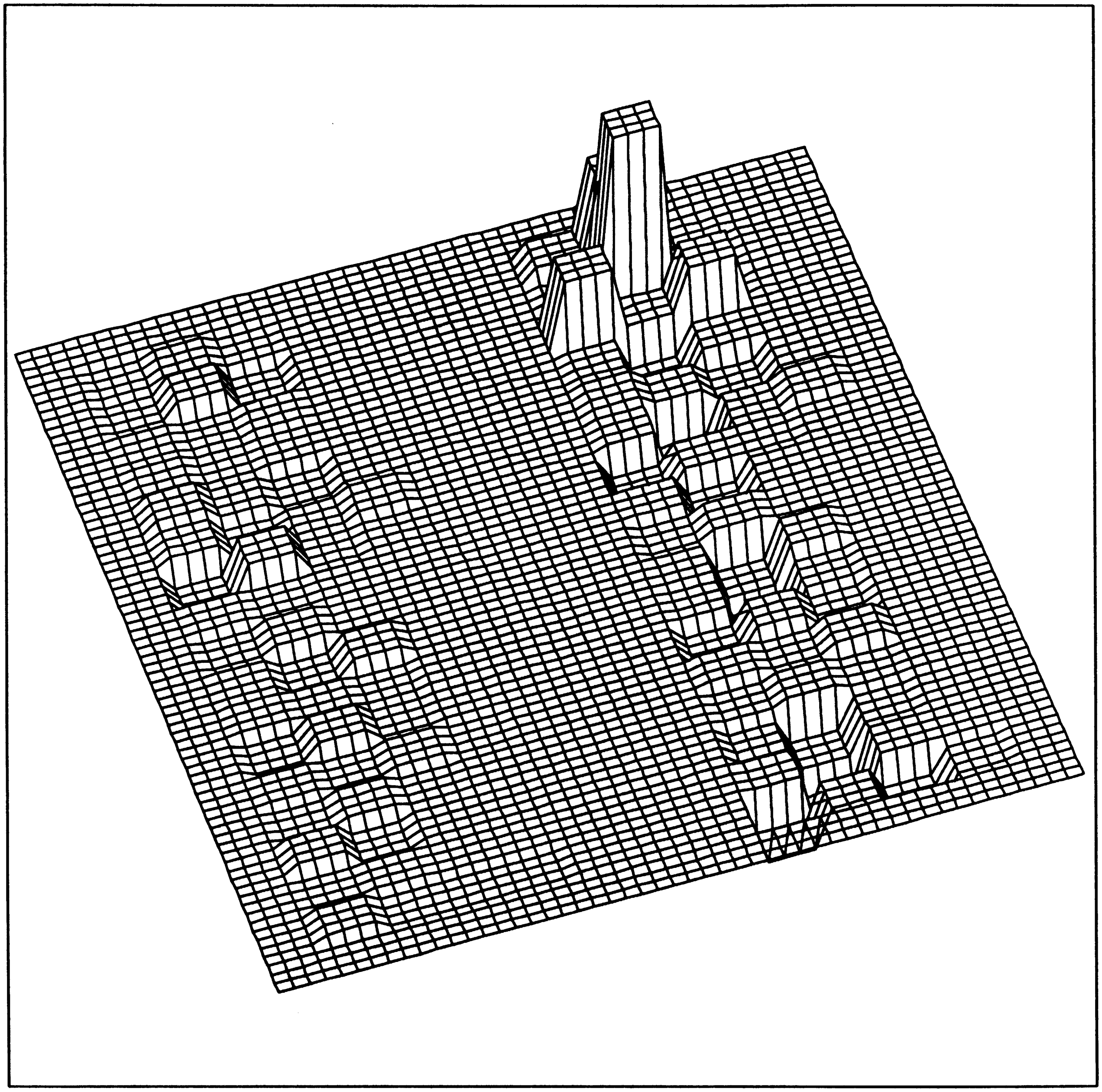


Figure 2: Residuals from Tables of Gross Omissions by Districts and Race
With and Without District Effects: Detago

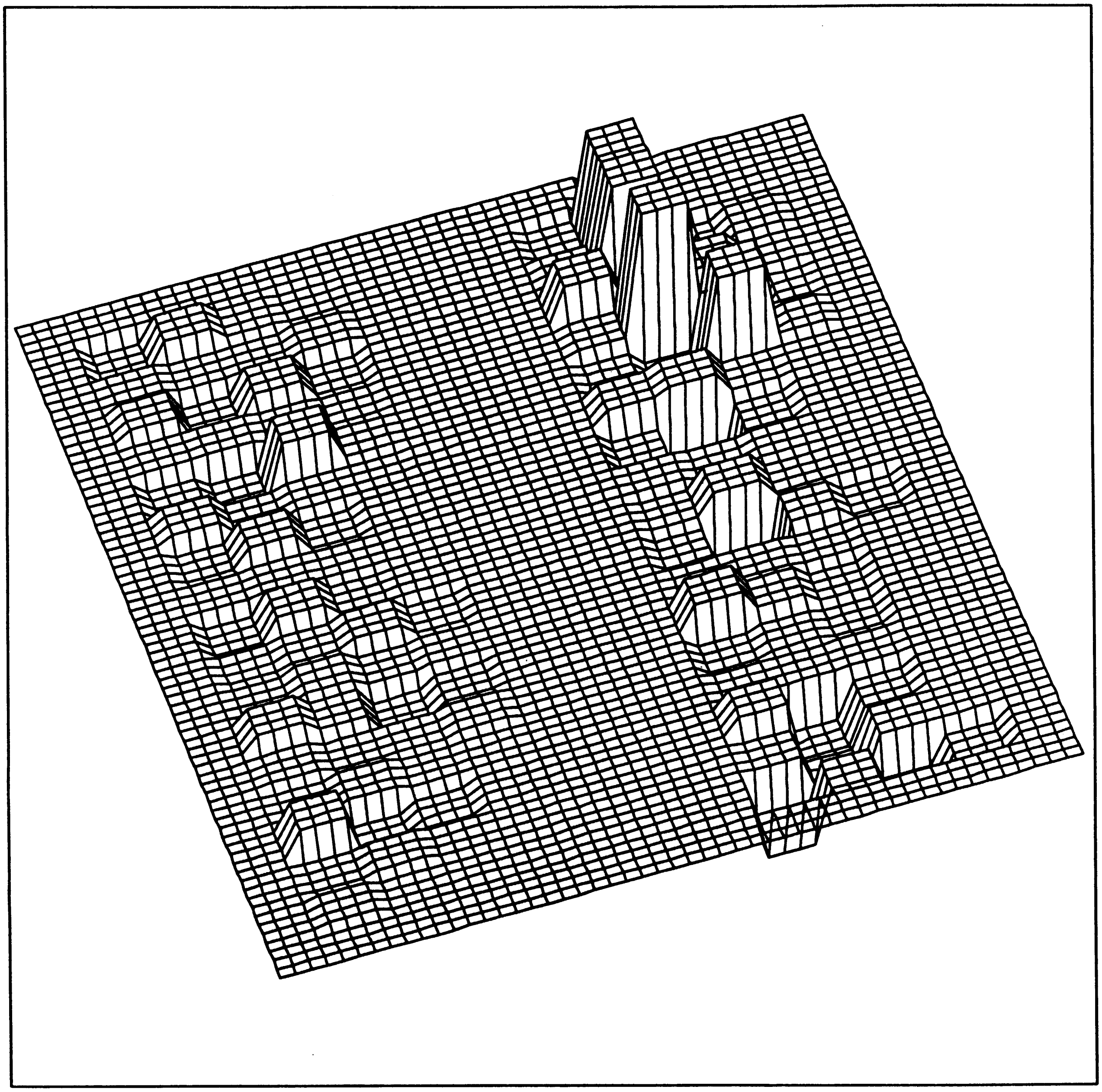


Figure 3: Residuals from Tables of Erroneous Enumeratons by Districts and Race
With and Without District Effects: Detago

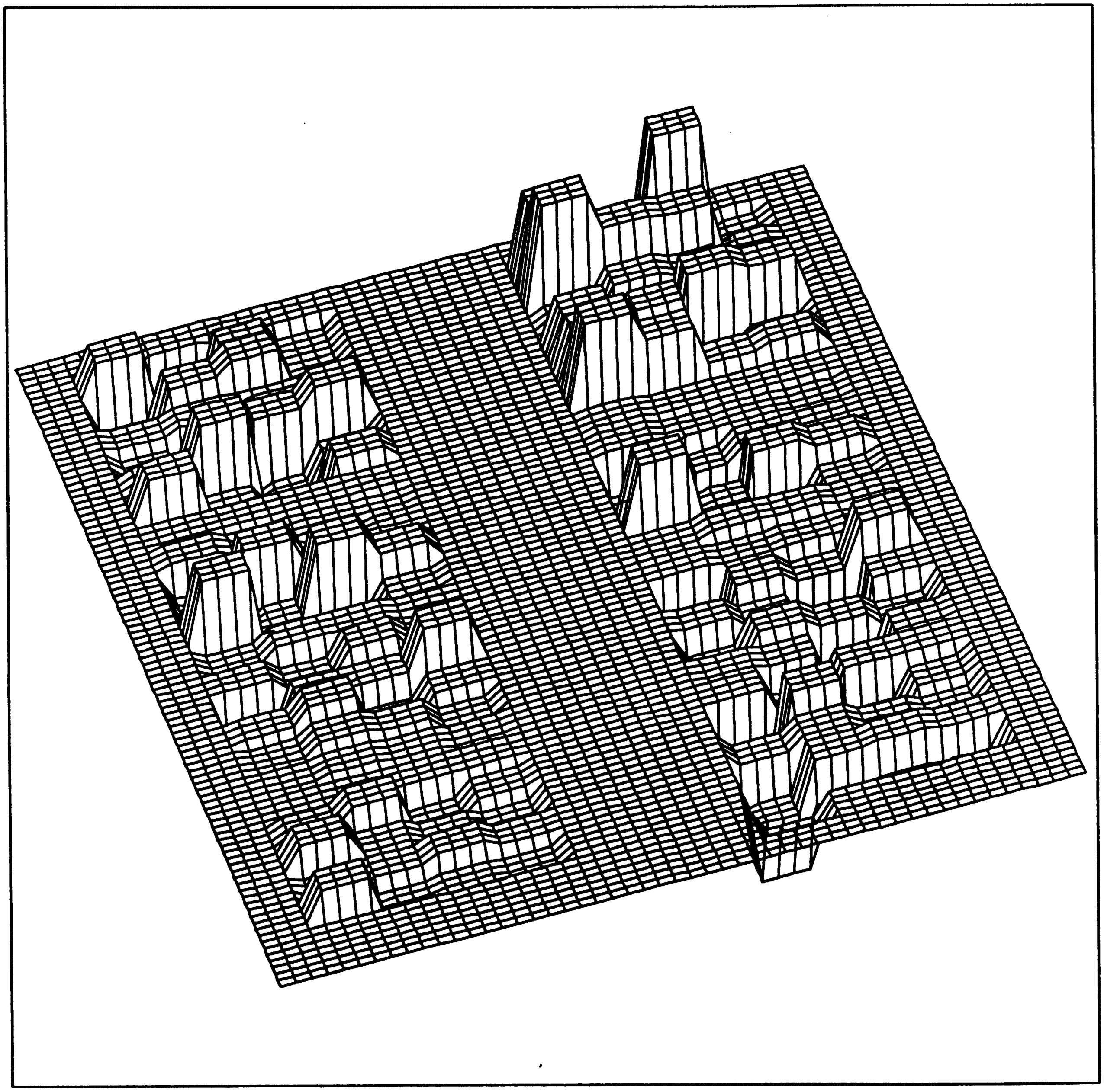


Figure 4: Residuals from Tables of Gross Omissions by Districts and Race
With and Without District Effects: New York

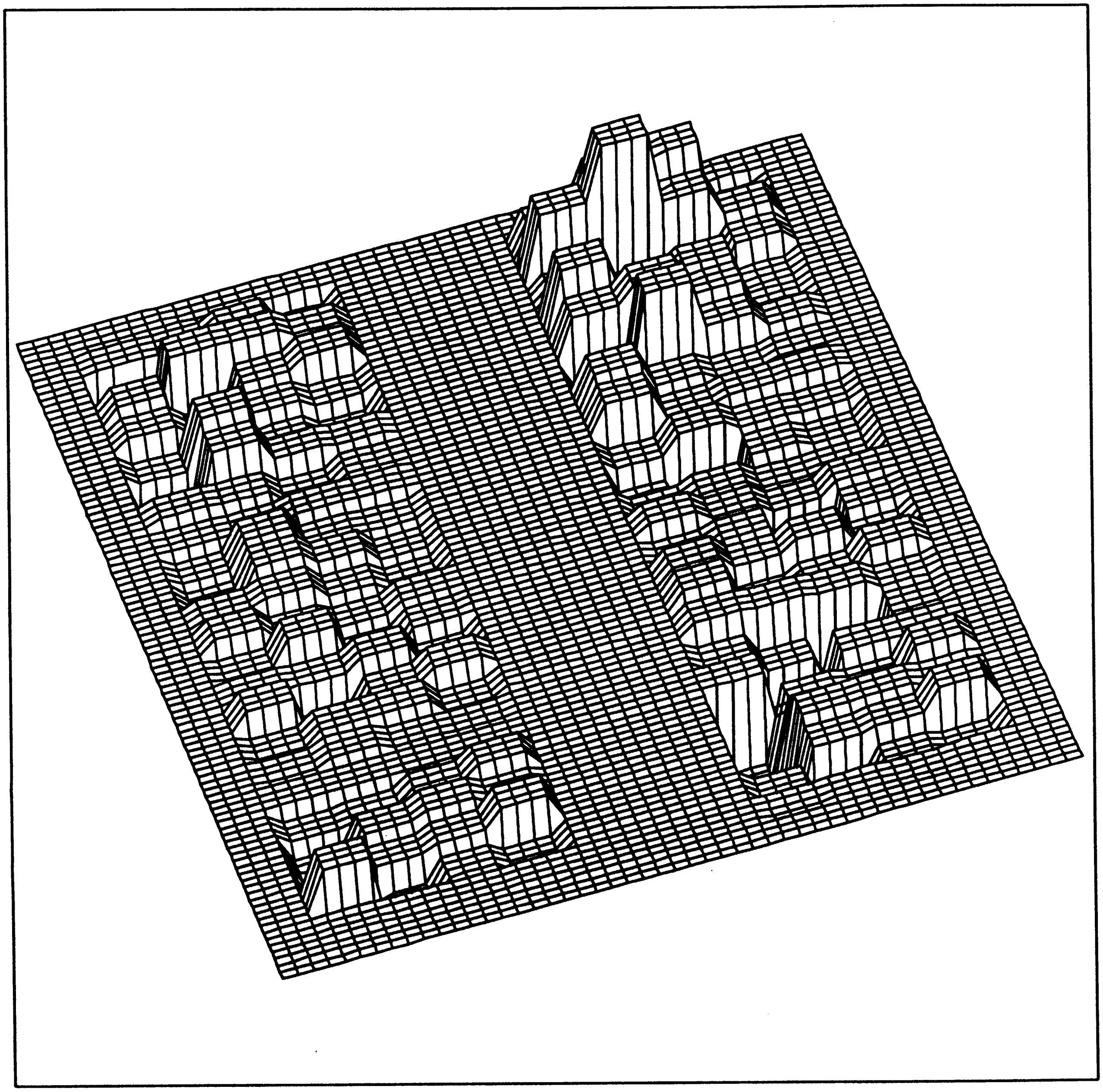


Figure 5: Residuals from Tables of Erroneous Enumeratons by Districts and Race With and Without District Effects: New York

III. Generating Hypothetical True Counts

In Section II, we discussed our strategy for generating Census counts which required both a table of hypothetical true counts and estimates of go(b,d) and ee(b,d), the probabilities that a given person living in district b and having a particular set of demographic characteristics d is either a gross omission or an erroneous enumeration. Whether a person represented in our table of hypothetical true counts appears once, twice or not at all in the Census is determined at random by a process equivalent to coin tossing using our estimates of go(b,d) and ee(b,d). Therefore, no matter what the distribution of people in our tables, each Census we generate will exhibit roughly the correct pattern of gross omission and erroneous enumeration rates.

Within each demographic group, these patterns should be captured by the Census Bureau's demoid-specific adjustment factors. There are certain ways of distributing people in our tables of hypothetical true counts, however, that are easier for adjustment to handle than others. For example, if the number of people in each demographic group was quite large and if each district contained all the people from only one such group, then no matter what estimates of go(b,d) or ee(b,d) we used, the adjustment procedure would perform better than if the each district contained only a small number of people scattered thinly between a number of different demoids.

Because we want our simulations to be a fair test of the challenges adjustment actually faces, we want the concentration of people in our tables to be broadly realistic. Therefore, when simulating Census results from PES districts, it is natural to construct tables of hypothetical true counts from the Census counts in our PES district chunks. In the final step of our simulations, however, we must apply a set of adjustment factors to generated Census counts from non-PES blocks. The natural choice for our tables of hypothetical true counts in this context would be the Census counts from a series of non-PES districts. Unfortunately, we do not have access to this data, and must again rely on the Census counts from our PES district chunks to construct our tables. While we have acknowledged in Section I. that our PES district chunks are different from Census districts, from the standpoint of our simulations we are mainly concerned with our tables broadly reproducing realistic patterns of residential segregation within districts by demoids. To summarize, then, the CAT simulations will generate Census results in both PES and non-PES districts from tables of hypothetical true counts built from the Census counts in our PES district chunks.

Just as certain patterns of population in our tables of hypothetical true counts were easier for adjustment to handle than others, the success of the adjustment procedure in carrying adjustment to non-PES districts depends in part on the similarity between the PES and non-PES districts. For example, adjustment will be much more effective when the PES and non-PES districts are exact copies of each other, than when their populations are concentrated in completely different demographic groups. Therefore, to obtain a fair appraisal of the adjustment process, we want the distribution of people by demoid in our tables for PES and non-PES districts to correspond only roughly, reflecting the diversity with which Census adjustment has to cope. To achieve this, for a fixed site we will build tables of hypothetical true counts for both PES and non-PES districts by sampling from a large pool of our PES district chunks, stratified by the percent of minority residents they contain, so that the resulting tables match the given site's minority profile.

As with the previous CAT simulations, we will build our pool by dividing all of the district chunks from each of sites 1 (New York), 2 (Detroit), 3 (Chicago), 4 (Arlington) and 7 (Berkeley) into five groups or bins based on their percentage of minority residents (referred to here as non-Whites): Bin 1 consists of all districts in our five sites with fewer than 20% minority residents, bin 2 consists of all districts in our five sites with between

20% and 40% minority residents, and so on. Figure 6 presents clearly how the 23 district offices from New York City are divided into the five bins. Each district office in New York is represented by one of the 23 boxes at the top of the figure, which are shaded according to the percentage of non-Whites living in the given district. A white box translates to a district which contains 0% minorities, while a black box represents a district composed entirely of minorities). For example, districts 9, 10 and 11 New York are made up of 14%, 65% and 86% non-whites, respectively. The arrows indicate to which bin of the Pool these districts will be assigned.

This process is repeated for each of the five sites under consideration, until we arrive at pool of 115 district offices. The numbers of offices per bin are given below.

	0-20%	20-40%	40-60%	60-80%	80-100%
n	37	18	19	16	25

There is nothing special about our district-level data with regard to this pooling operation. When we extend our simulations to block-level data, the pool is constructed similarly, only this time the individual block clusters from each site are divided up into five bins. When this is done, we end up with a total of 546 block clusters in our pool with the following distribution.

	0-20%	20-40%	40-60%	60-80%	80-100%
n	170	70	30	48	228

As noted above, we will construct our district-by-demoid tables of hypothetical true counts for a given site by sampling from this pool so that the resulting tables match the site's minority profile. Figure 7 presents a schematic representation of how this sampling is performed for New York City. Again, each of the 23 districts in New York is represented by a box at the top of the figure. Here, each box is labeled with the number of the bin in the pool to which the original district belonged. Recall that districts 9, 10, and 11 belonged to bins 1, 4, and 5, respectively. Thus, when building tables of hypothetical true counts for this site from the pool, we will randomly select a district chunk from bin 1 to represent district 9, and take the census counts (by demoid) from the selected chunk as the true population count (by demoid) for district 9 in our table. Similarly, we select a district chunk from bin 4 to represent district 10, and take the census counts from the selected district chunk as true population counts for district 10 in our table. We repeat this process for all 23 districts in New York. Again, the process will be the same whether we are creating a table of hypothetical true counts for PES or non-PES districts for New York City.

This scheme for generating district-by-demoid tables of hypothetical true counts as described in Figure 7 applies just as easily to any one of our other districts, and is the same when we eventually consider block-level data. As we will see in the next Section, this process will play a very important role in the CAT simulations. Essentially, we will depend on these tables to provide a measure against we will judge the effectiveness of Census adjustment.

Figure 6: Creation of Pool from District Chunks in New York Site

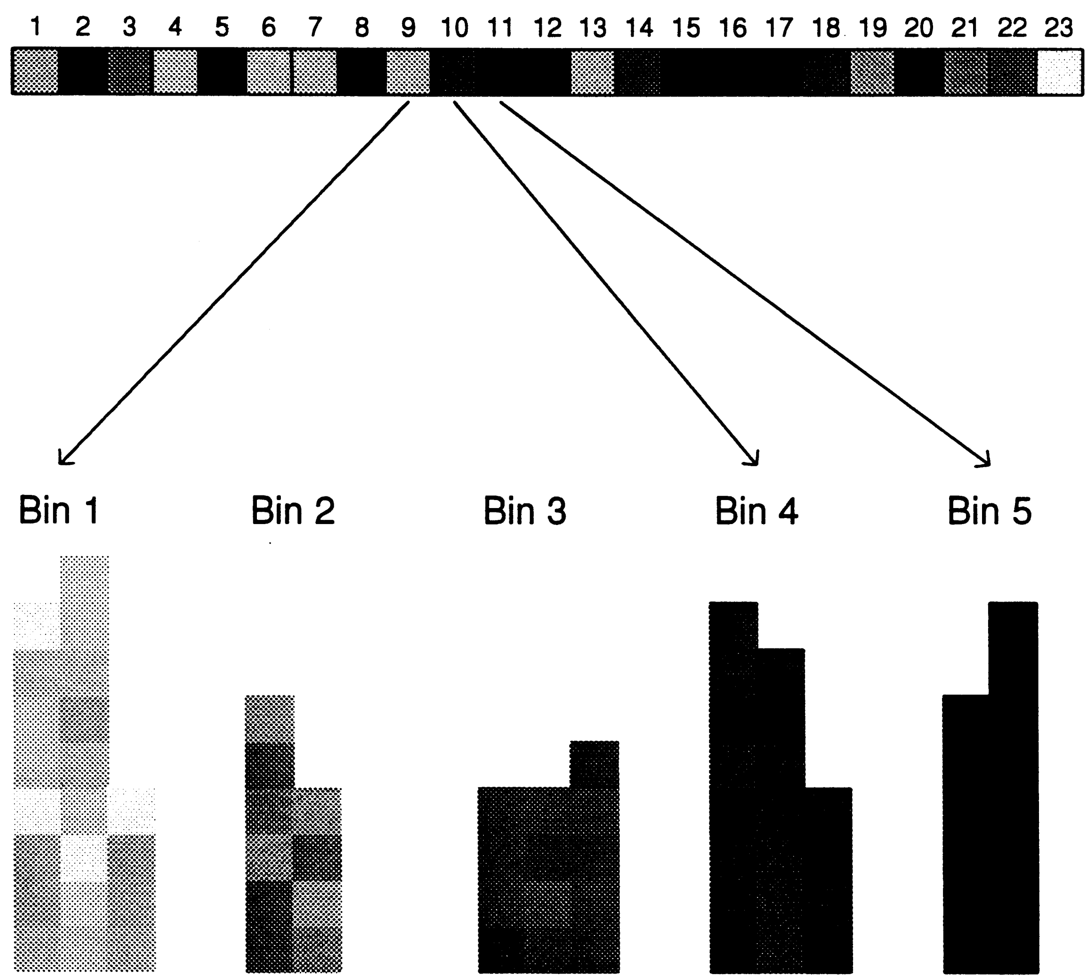
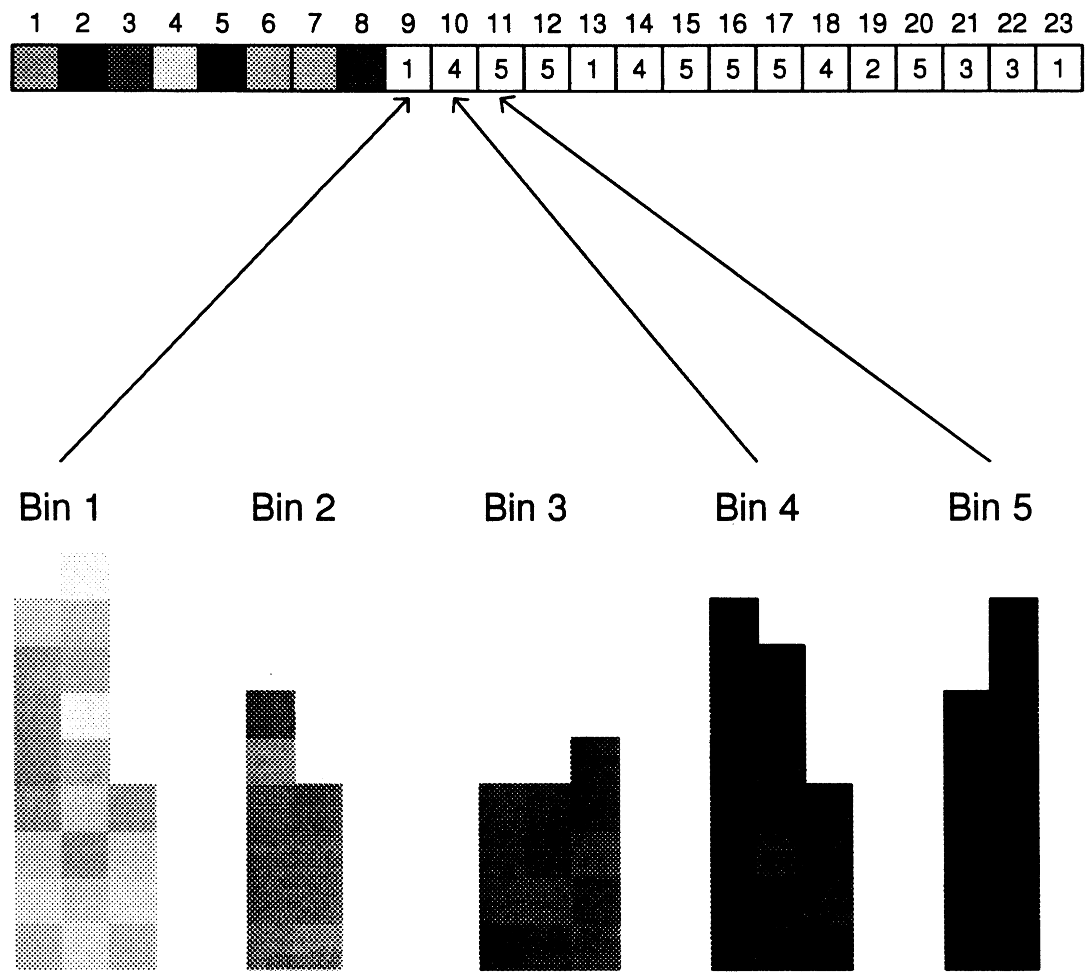


Figure 7: Sampling from the Pool, District Chunks in New York Site



IV. Design of CAT Simulations

The CAT simulations have changed considerably since their introduction by Kenneth Wachter in his report from June of 1991. We have already seen changes in Section II, when we introduced FIT1 through FIT4 for estimating go(b,d) and ee(b,d) rather than fitting log-linear models to tables of go- and ee-rates. Further changes will be introduced in this Section as we present the current design of the CAT simulations. To simplify our discussion, we will concentrate on applying the simulations to New York, but will present a general algorithm at the end of the Section.

In what follows, we will divide the adjustment process into two stages, and consider separately the way in which they are treated by the CAT simulations. We take as the first stage of adjustment the calculation of demoid-specific adjustment factors from Census and PES counts on a set of PES districts. In the second stage, these adjustment factors are applied to Census counts in non-PES districts.

1. **PES Districts.** To start, we create a district-by-demoid table of hypothetical true counts by sampling from our pool of district chunks as described in Figure 7 of Section III. The population in this table will represent the true counts in a set of PES districts. In addition to labeling each person in this table as being counted once, twice or not at all by the Census, we also have to decide whether or not he or she has been included

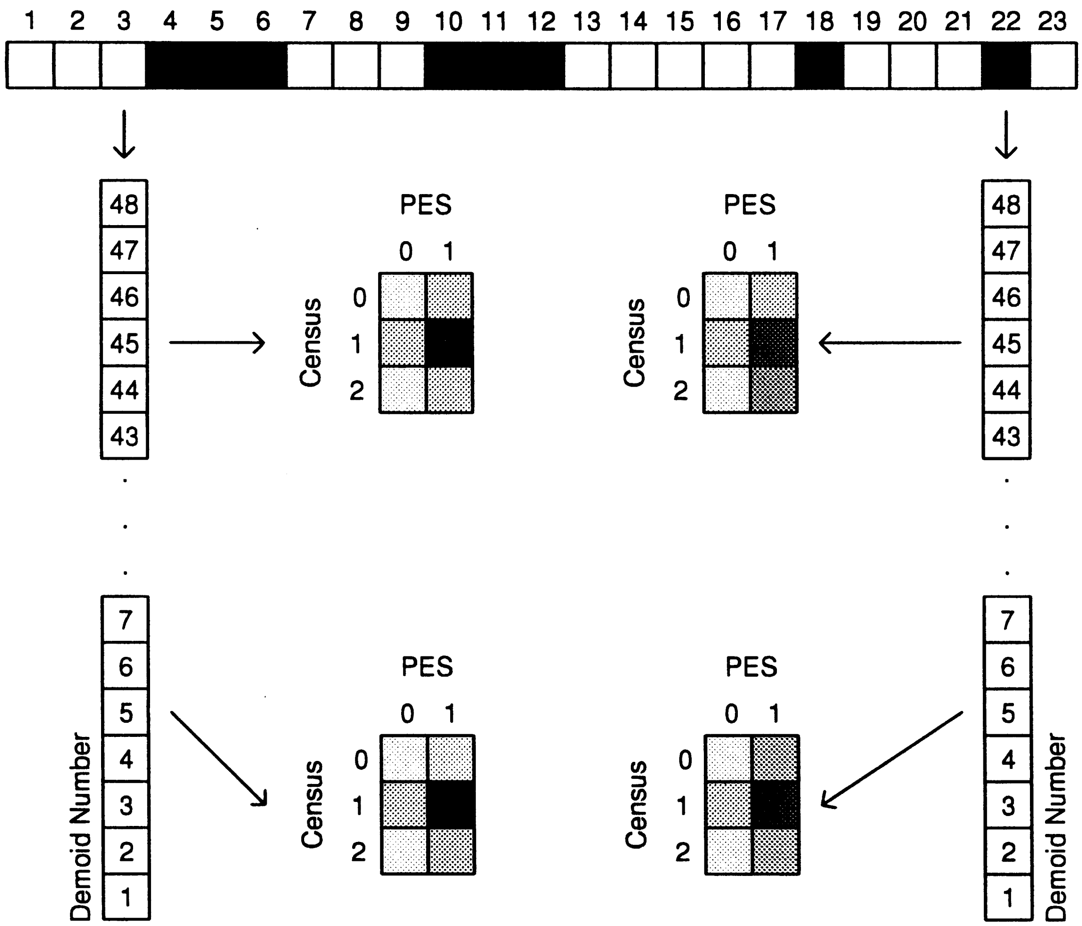
in the PES. For the purpose of this simulation, we will assume that each person in our table has a 90% chance of being in the PES, independent of their status in the Census. The six possible outcomes and their probabilities are given in the table below.

		PES 0	PES 1
Census	0	$0.1 \times go(b,d)$	$0.9 \times go(b,d)$
Census	1	$0.1 \times (1 - go(b,d) - ee(b,d))$	$0.9 \times (1 - go(b,d) - ee(b,d))$
Census	2	$0.1 \times ee(b,d)$	$0.9 \times ee(b,d)$

Table 6

Figure 8 presents the results of applying FIT3 to our PES data from New York City. The boxes along the top edge again represent the 23 districts in the site, and are shaded so that a black box indicates a heavy district while a white box indicates a light district. We have included tables like the one presented above for districts 3 and 22, and demoids 5 (White males between the ages of 45 and 64) and 45 (Asian females between the ages of 20 and 29). In each case, the cells are shaded according to the probabilities of the outcomes they denote. From this fit we find that each 47 year old While male living in district 3 would have a 0.4% chance of being missed by both the Census and the PES, and a 10.8% chance of being caught by the PES, but counted twice by the Census, while each 25 year old Asian female living in district 22 has a 6.5% chance of being missed by the PES but correctly enumerated in the Census. We emphasize that each group of people in New York, classified by district

Figure 8: Combination of Estimated Probabilities for New York Site



and demographic characteristics, will have different probabilities associated with each of the six possible Census and PES outcomes.

Using these probabilities, we randomly assign each person in our table of hypothetical true counts to one of the six cells in Table 6. We then aggregate the resulting Census and PES counts by demoid across districts to form the Census Bureau's demoid-specific adjustment factors. In the second stage of the adjustment process, we apply these adjustment factors to a set of non-PES districts.

2. **Non-PES Districts** Once again, we form a table of hypothetical true counts by sampling a set of 23 district chunks from our pool as explained in the Figure 7 of Section III. This demoid-by-district table of counts will represent the true population for a set of non-PES districts. We simulate Census counts from these districts by labeling each person in the table as being counted once, twice or not at all according to our estimates of go(b,d) and ee(b,d) (see Table 5 in Section II). Finally, we apply the adjustment factors calculated in part 1. from our PES districts to these simulated Census counts. We can now compare both Census and adjusted counts with our table of hypothetical true counts and assess the improvement offered by the Bureau's adjustment procedure.

   Clearly, however, our results will be different if we apply these adjustment factors to another random Census realization from the same set of non-PES districts. Therefore, in our simulations, for each district-by-demoid table of counts representing a set of non-PES districts, we will generate several sets of random Census counts, using the same estimates of go(b,d) and ee(b,d), and apply the same adjustment factors to each. This will give us an idea of the average behavior of a fixed set of adjustment factors on a given set of non-PES districts.

   Generating several random Census realizations for each fixed table of hypothetical true counts represents only one form of randomness in our simulations, however. If, for example, we select a different set of chunks from the pool to form our table of hypothetical true counts representing the set of non-PES districts, we would observe different levels of improvement from the adjustment procedure. Therefore, for a given set of PES districts and their associated adjustment factors, we will sample several different sets of non-PES districts to get an idea for the average behavior of these adjustment factors on a variety of non-PES districts, each having roughly the same pattern of residential segregation by demoids as the original PES districts.

To summarize, for a fixed site, the CAT simulations sample sets of of district chunks to represent PES and non-PES districts. A Post-Enumeration Survey is simulated in the PES districts while a Census is simulated in both. The PES and Census counts from the PES districts are used to calculate adjustment factors that are applied to the non-PES districts. We judge the average improvement of the adjustment process by sampling several sets on non-PES districts for each set of PES districts, and by generating several random Census realizations from each non-PES district.

For a fixed site, the results presented in Section V were obtained by sampling a total of ten sets of PES districts from our pool of district chunks. For each such set, we sampled five sets of non-PES districts, each of which were used to generate five random Census realizations from estimates of go(b,d) and ee(b,d). In all, we are left with 250 Census-Adjustment-Truth comparisons per site from which we obtain a fair appraisal of the success of the adjustment procedure. This process was repeated for estimates of go(b,d) and ee(b,d) derived from each of FIT1 and FIT3.

V. CAT Simulation Results

Complete district- and block-level CAT simulations have been run for New York, Arlington, Berkeley and Detago (the combined PES site of Detroit and Chicago) according to the plan presented at the end of Section IV. That is, for a fixed site, we sample a total of ten sets of PES districts from our pool of district chunks. For each such set, we sampled five sets of non-PES districts, each of which were used to generate five random Census realizations. In all, we are left with 250 Census-Adjustment-Truth comparisons per site.

To judge the performance of the adjustment procedure, we calculate two measures of improvement for each of these 250 sets of counts. For the first, we aggregate our Census, adjusted and hypothetical true counts by district, and report the proportion of districts in a given site for which the adjusted count is closer to the truth than the Census count. Simply looking at district-by-district totals, however, is not sufficient to adequately judge adjustment's performance. Because the adjustment process changes the total population for a site, we will also be concerned with how adjustment changes a given district's share of the population in the site (expressed as the proportion of the district's population relative to the population for the whole site). Our second measure of improvement, therefore, will be the proportion of districts that have their share of the total population for the site brought closer to their true share than that calculated using Census counts.

As noted previously, we have completed simulation runs using estimates of go(b,d) and ee(b,d) from FIT1 and FIT3. Recall that FIT1 estimates the probabilities of gross omission or erroneous enumeration under the Census Bureau's assumption of homogeneity across districts. Results from simulations run using these estimates of go(b,d) and ee(b,d) can therefore be viewed as an appraisal of how Census adjustment performs under the best possible circumstances. Estimates of go(b,d) and ee(b,d) derived from FIT3, however, include both district effects and a district by demoid interaction as explained in Section II. Results from simulations run using these estimates can be viewed as describing the behavior of Census adjustment under more broadly realistic conditions.

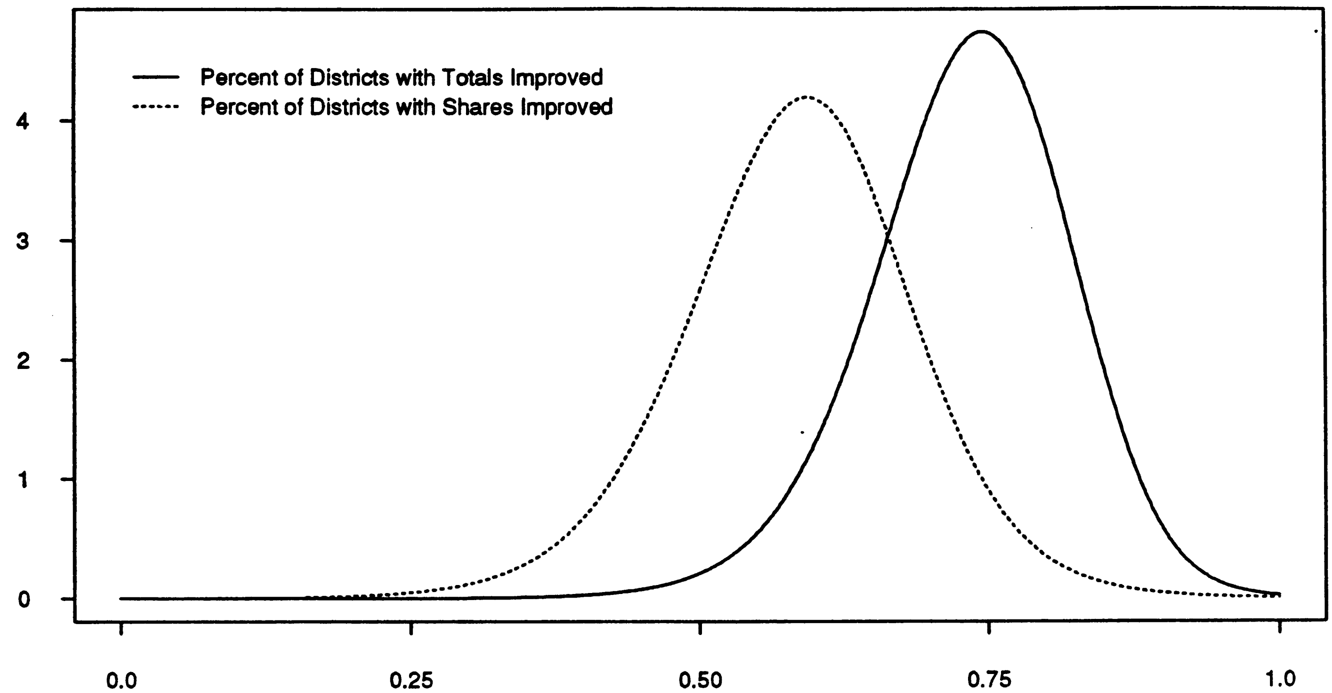
Consider first our district-level simulations. For a fixed site and a fixed set of estimates of go(b,d) and ee(b,d), we have 250 Census-Adjustment-Truth comparisons. For each, we know what proportion of the districts in the site had their totals brought closer to the truth, and what proportion had their shares brought closer to the truth by the adjustment process. The distribution of these quantities across the 250 simulations is given by a pair of histograms. Take for example Figure 9 in the appendix. We present here the simulation results from Detago using estimates of the probability of gross omission and erroneous enumeration derived from FIT1. The upper histogram shows the distribution over 250 simulation runs of the proportion of the 17 districts in the site that had their totals improved by adjustment. The lower histogram shows the distribution over 250 simulations of the proportion of the 17 districts in the site that had their shares improved by adjustment. The format for the results for each of the remaining three sites is similar, and are presented in Figures 5 through 20 in the appendix.

To simplify interpretation somewhat, we also present "smoothed" versions of these histograms. To be more precise, we calculate the logspline density estimate of the distributions associated with each of our two measures of improvement over a given set of 250 simulation runs. These can now be plotted in the same graph and the results for totals and shares can be compared directly. Figure 13 presents such a plot for Detago (again using probability estimates from FIT1). The solid line is a smoothed version of upper histogram in Figure 9 of the appendix, while the dashed line is a smoothed version of the lower histogram in the same Figure. Futher discussion of logspline density estimation can be found in Kooperberg and Stone (1991,1992).

Figures 9 through 16 are similar to Figure 13 described above, but cover all combinations of sites and probability estimates (FIT1 or FIT3). Each of these plots refer to district-level simulations. Results for block-level runs are given in Figure 17 through 24.

The important statistics to be culled from these plots include the number of times (out of 250) that adjustment makes a majority of the district totals for a given site worse than the Census counts, as well as the number of times that adjustment makes a majority of the district shares for a given site worse than those calculated from Census counts. In addition, we are interested in the average (over 250 simulations) proportion of districts that have their totals or shares improved. These results, as well as similar statistics calculated for our block-level simulations are presented in Tables 7 and 8.

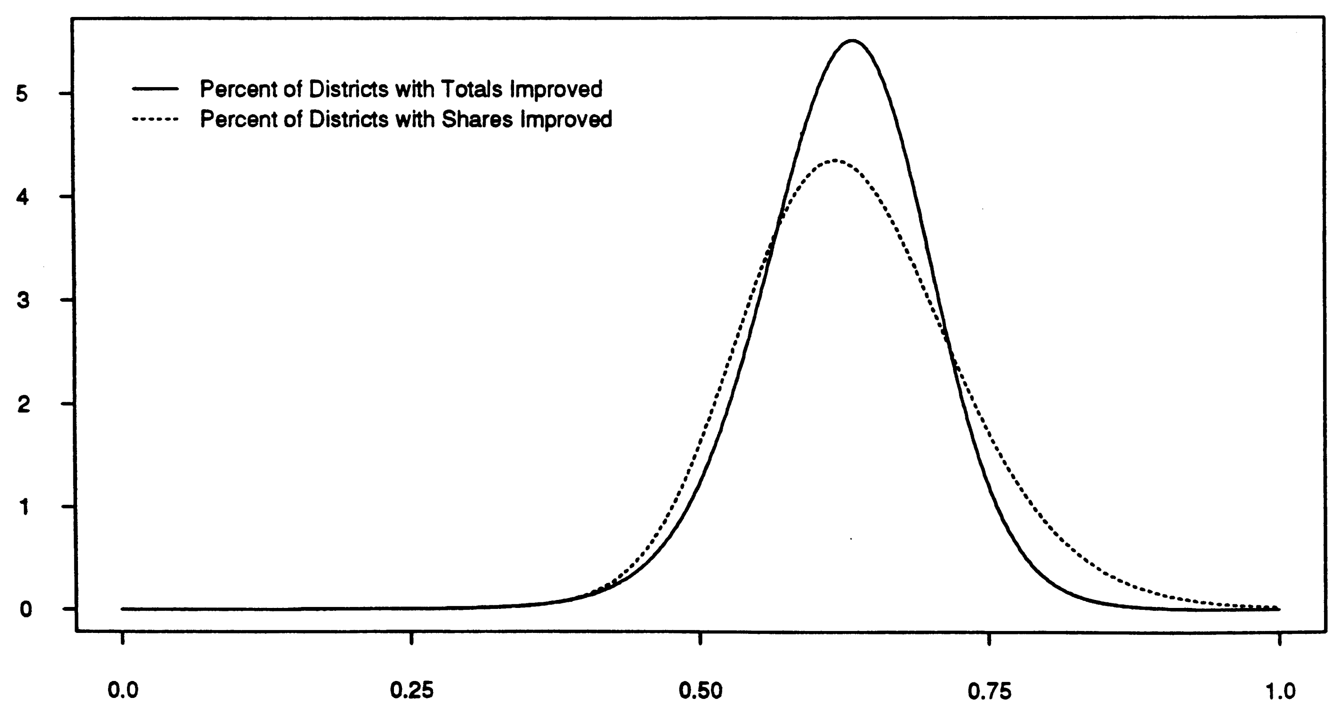
Figure 9: Percent of Districts Improved, New York, Without District Effects
(250 Simulations)



73% Mean Improvement of Totals, 58% Mean Improvement of Shares

1/250 had < 50% Improved Totals, 46/250 had < 50% Improved Shares

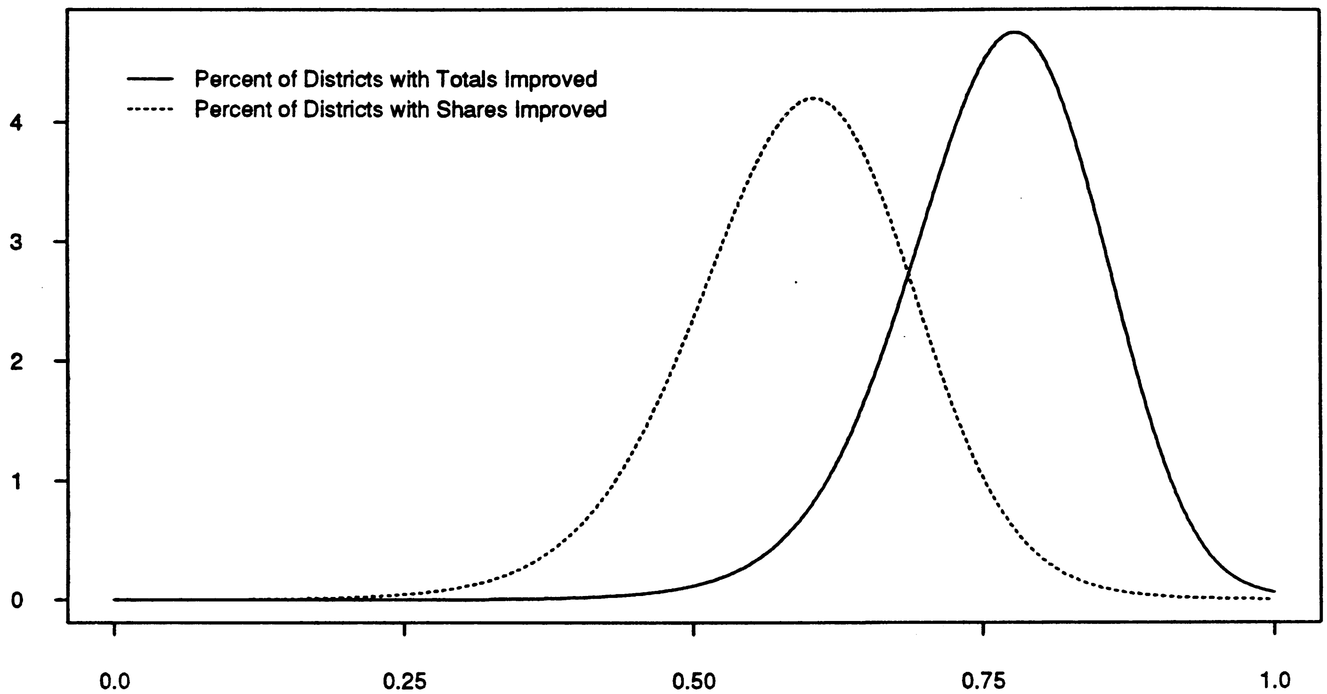
Figure 10: Percent of Districts Improved, New York, With Effects + Int
(250 Simulations)



63% Mean Improvement of Totals, 63% Mean Improvement of Shares

11/250 had < 50% Improved Totals, 21/250 had < 50% Improved Shares

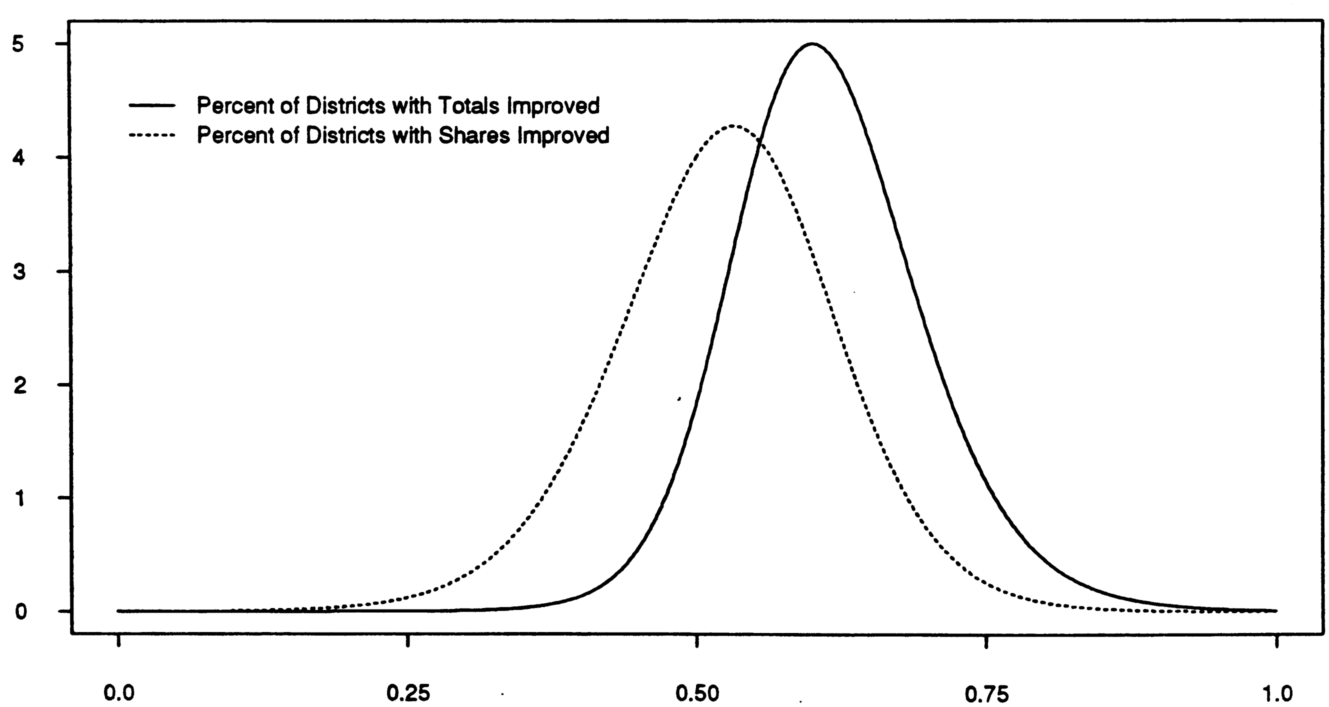
Figure 11: Percent of Districts Improved, Arlington, Without District Effects
(250 Simulations)



76% Mean Improvement of Totals, 59% Mean Improvement of Shares

0/250 had < 50% Improved Totals, 46/250 had < 50% Improved Shares

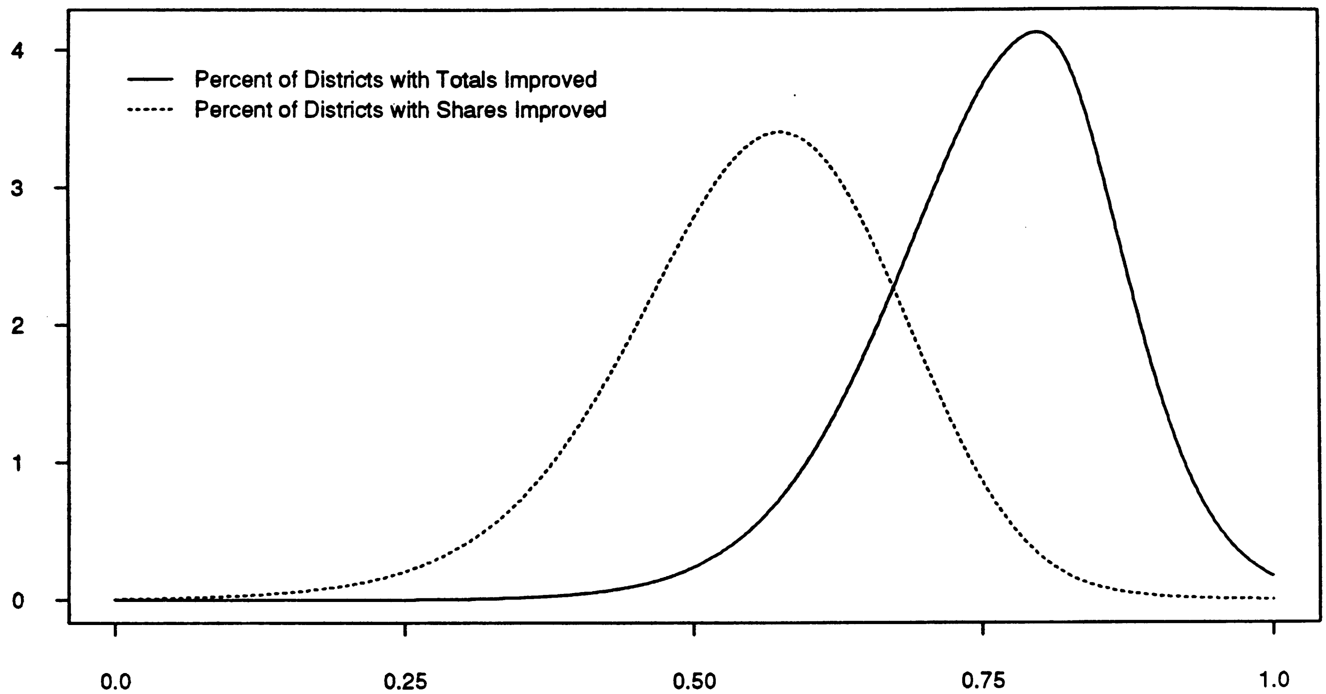
Figure 12: Percent of Districts Improved, Arlington, With Effects + Int
(250 Simulations)



62% Mean Improvement of Totals, 53% Mean Improvement of Shares

18/250 had < 50% Improved Totals, 102/250 had < 50% Improved Shares

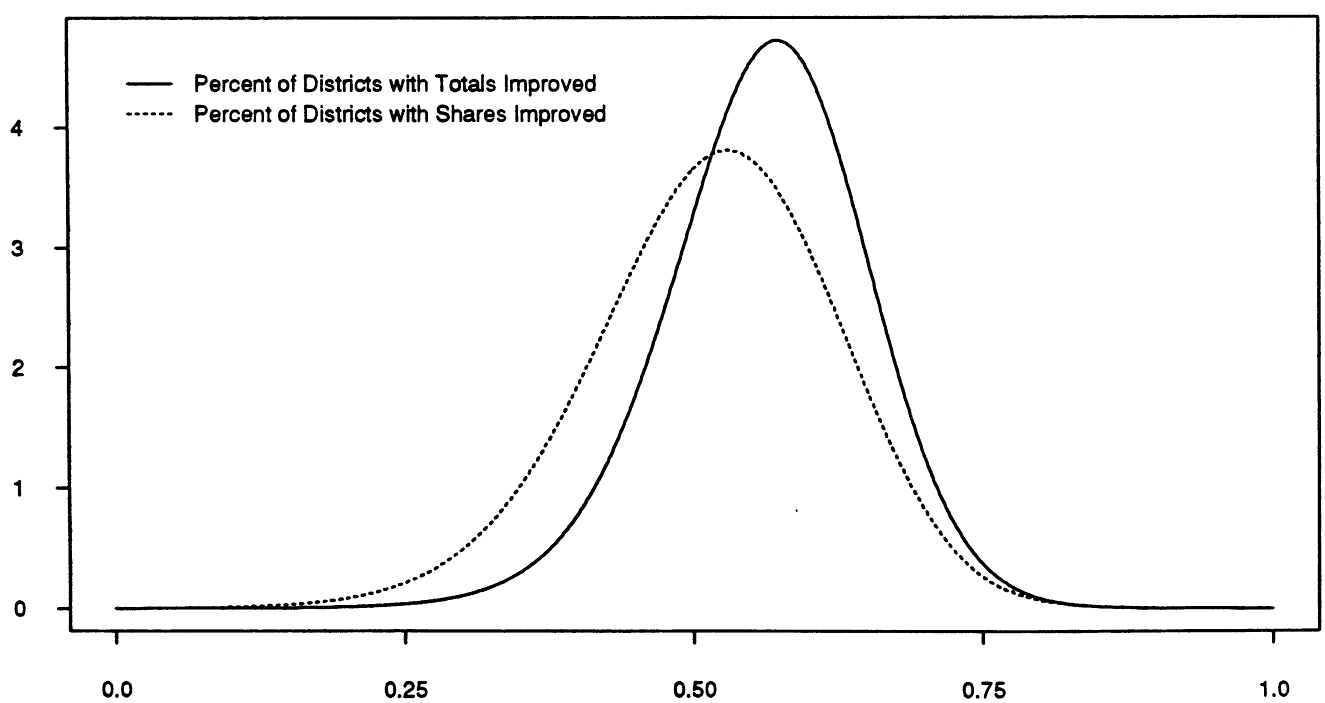
Figrure 13: Percent of Districts Improved, Detago, Without District Effects
(250 Simulations)



76% Mean Improvement of Totals, 55% Mean Improvement of Shares

3/250 had < 50% Improved Totals, 81/250 had < 50% Improved Shares

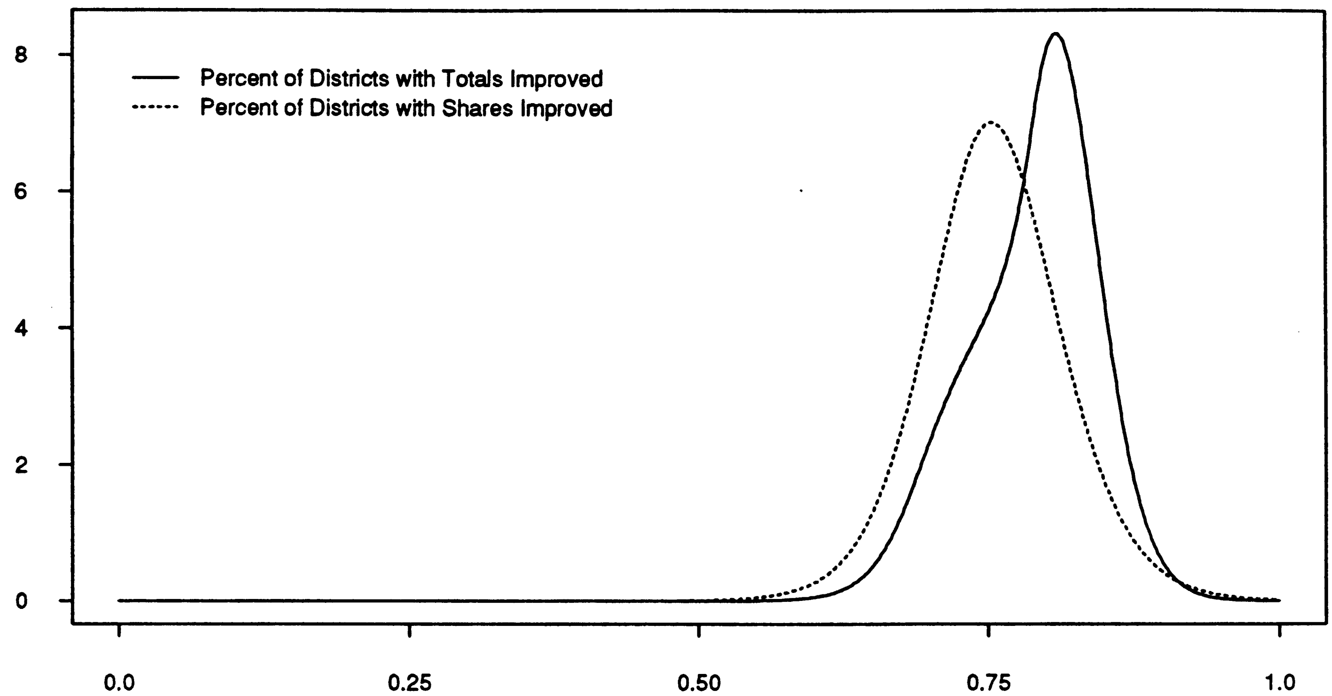
Figure 14: Percent of Districts Improved, Detago, With Effects + Int
(250 Simulations)



56% Mean Improvement of Totals, 52% Mean Improvement of Shares

59/250 had < 50% Improved Totals, 109/250 had < 50% Improved Shares

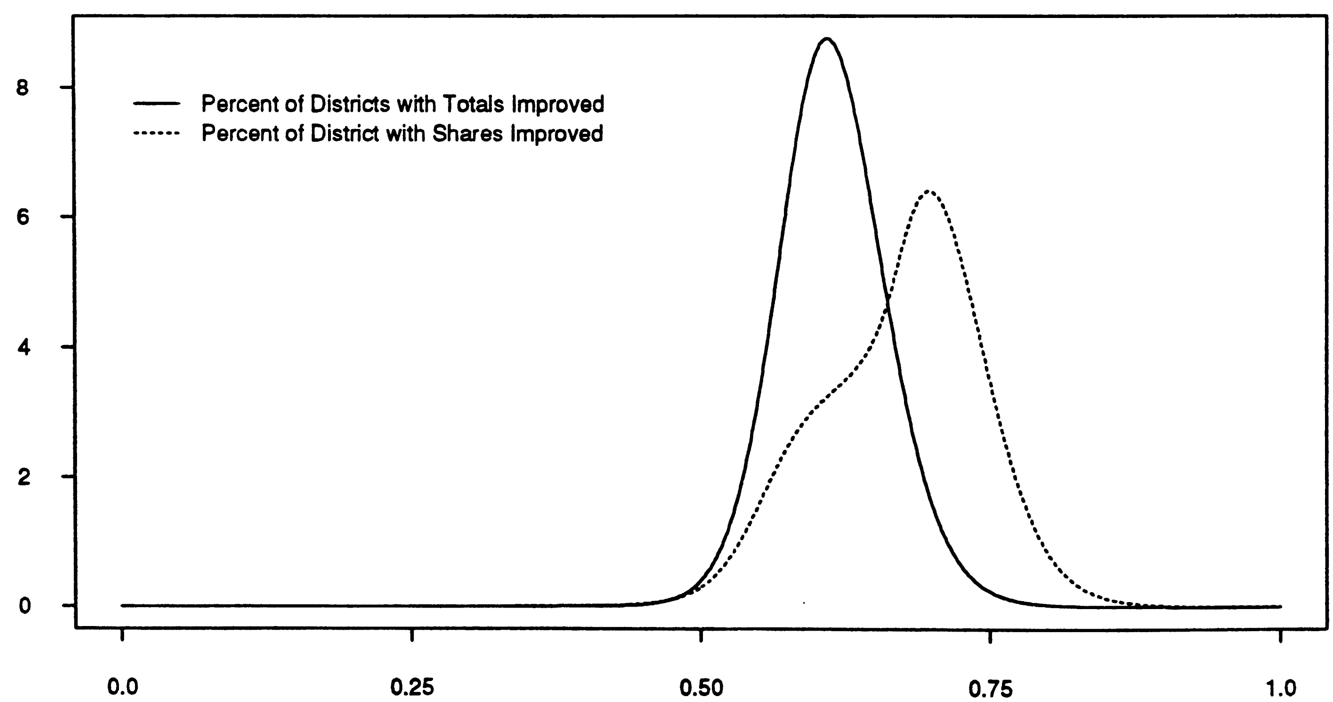
Figure 15: Percent of Districts Improved, Berkeley, Without District Effects
(250 Simulations)



79% Mean Improvement of Totals, 76% Mean Improvement of Shares

0/250 had < 50% Improved Totals, 0/250 had < 50% Improved Shares

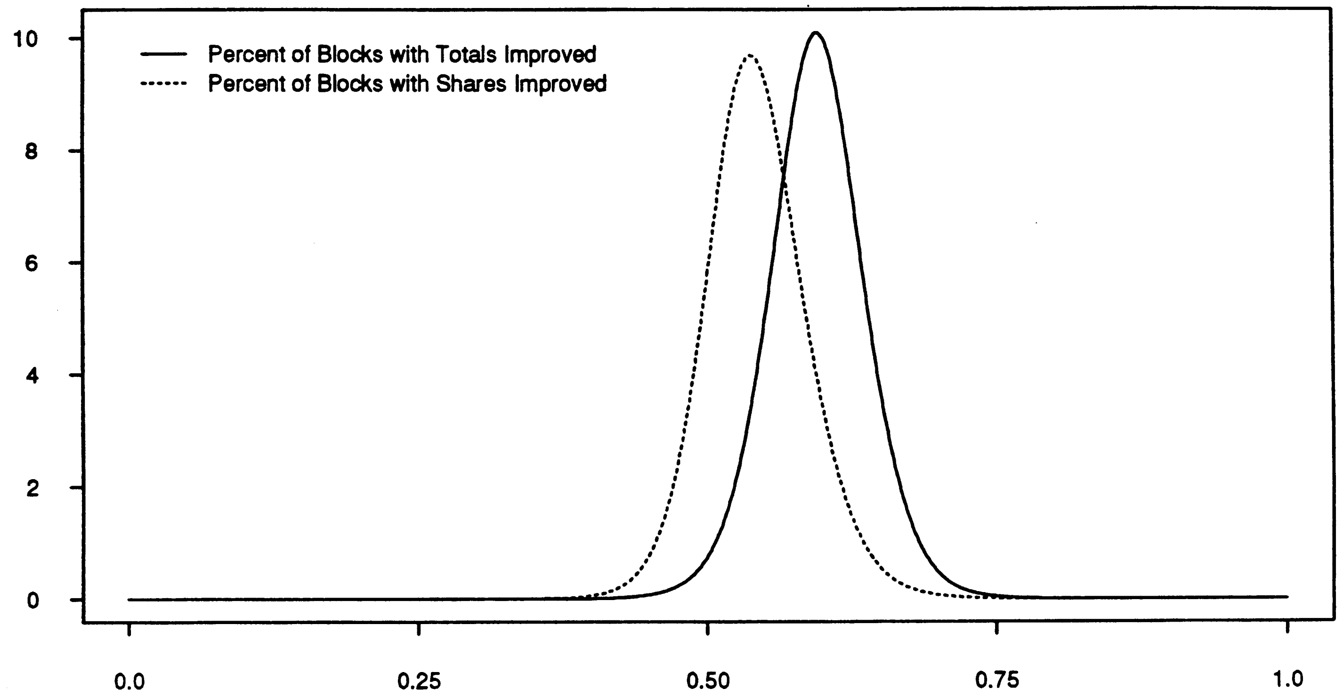
Figure 16: Percent of Districts Improved, Berkeley, With Effects + Int



62% Mean Improvement of Totals, 67% Mean Improvement of Shares

1/250 had < 50% Improved Totals, 1/250 had < 50% Improved Shares

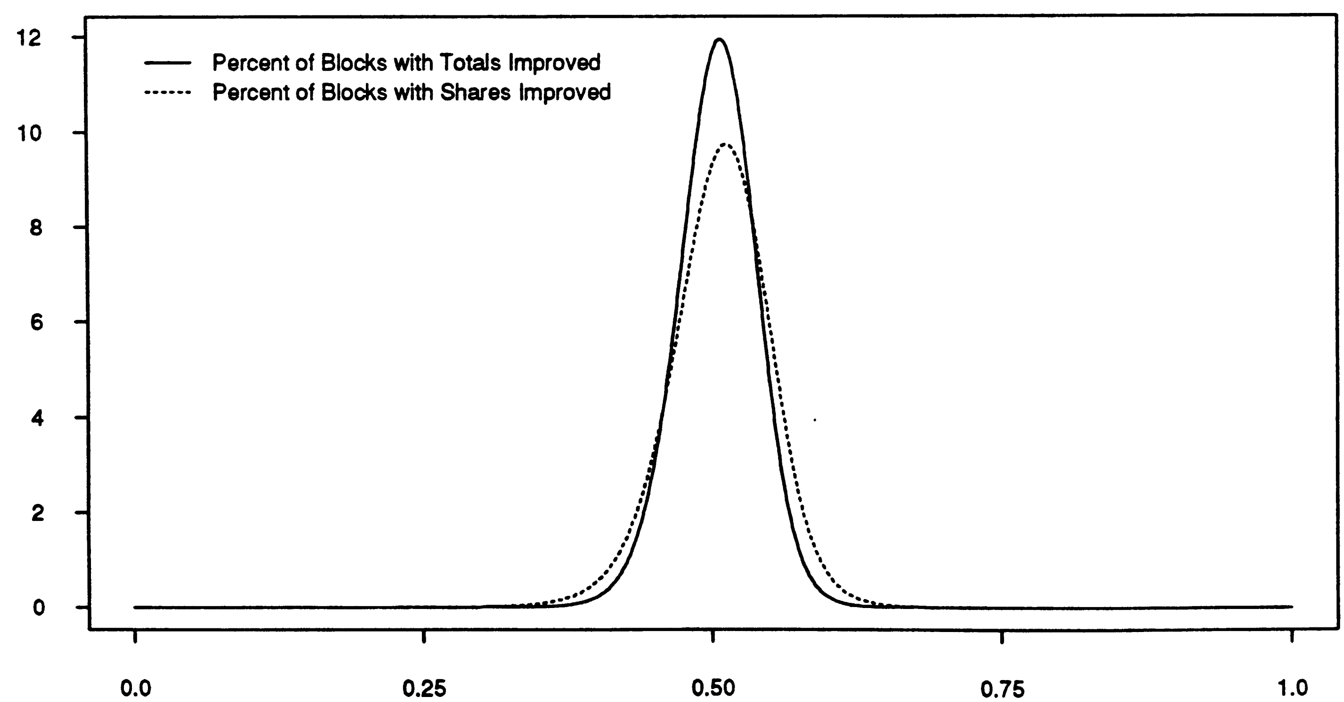
Figure 17: Percent of Blocks Improved, New York, Without Block Effects
(250 Simulations)



60% Mean Improvement of Totals, 54% Mean Improvement of Shares

5/250 had < 50% Improved Totals, 34/250 had < 50% Improved Shares

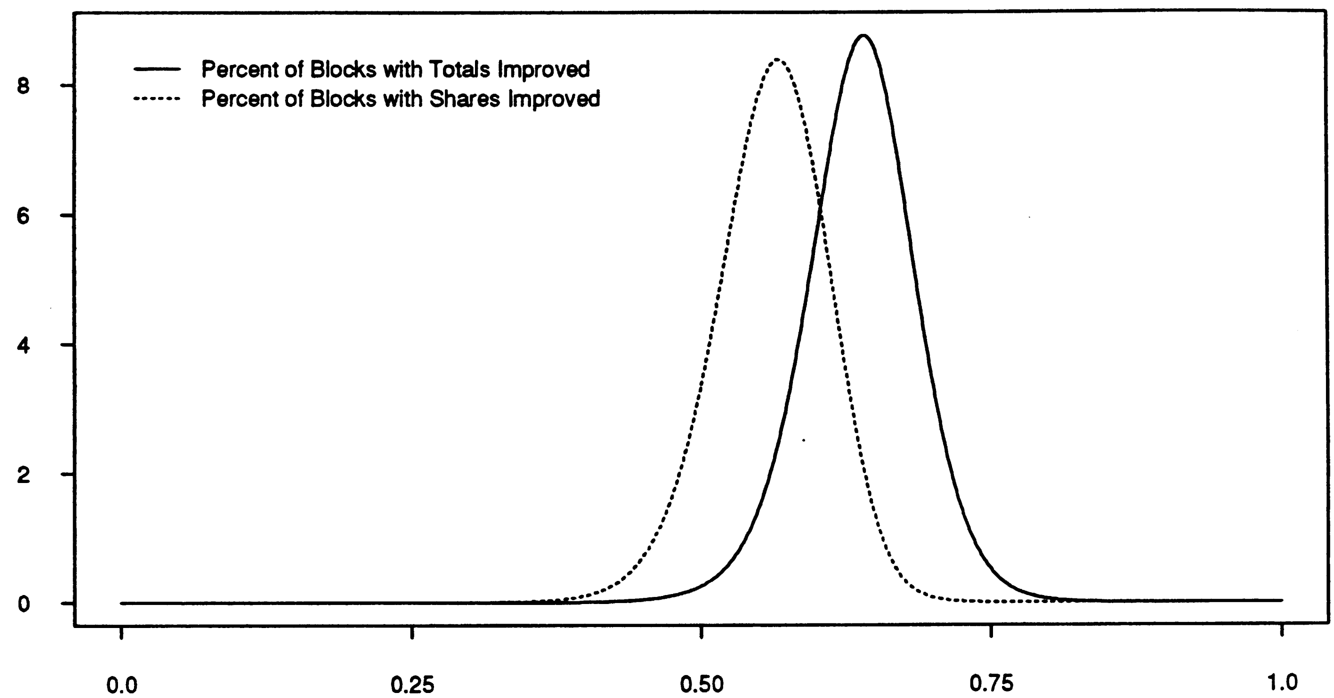
Figure 18: Percent of Blocks Improved, New York, With Effects + Int
(250 Simulations)



50% Mean Improvement of Totals, 51% Mean Improvement of Shares

108/250 had < 50% Improved Totals, 112/250 had < 50% Improved Shares

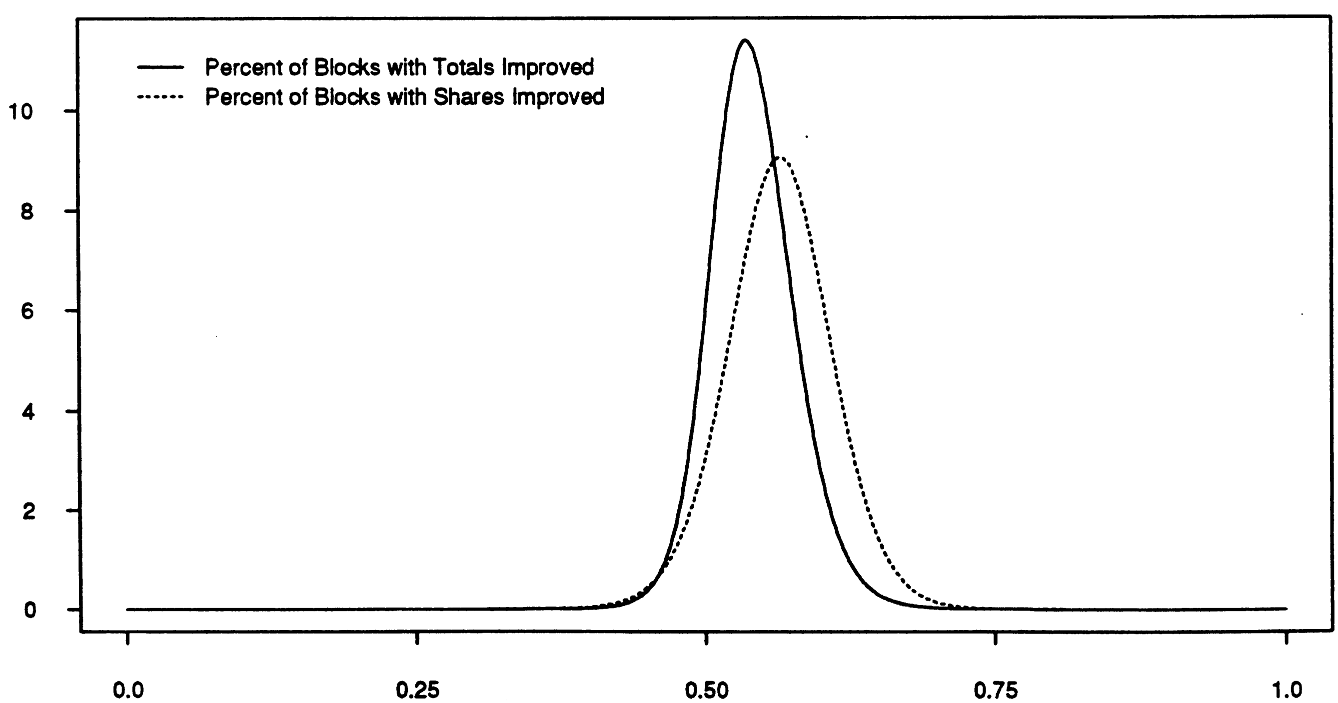
Figure 19: Percent of Blocks Improved, Arlington, Without Block Effects
(250 Simulations)



64% Mean Improvement of Totals, 56% Mean Improvement of Shares

0/250 had < 50% Improved Totals, 22/250 had < 50% Improved Shares

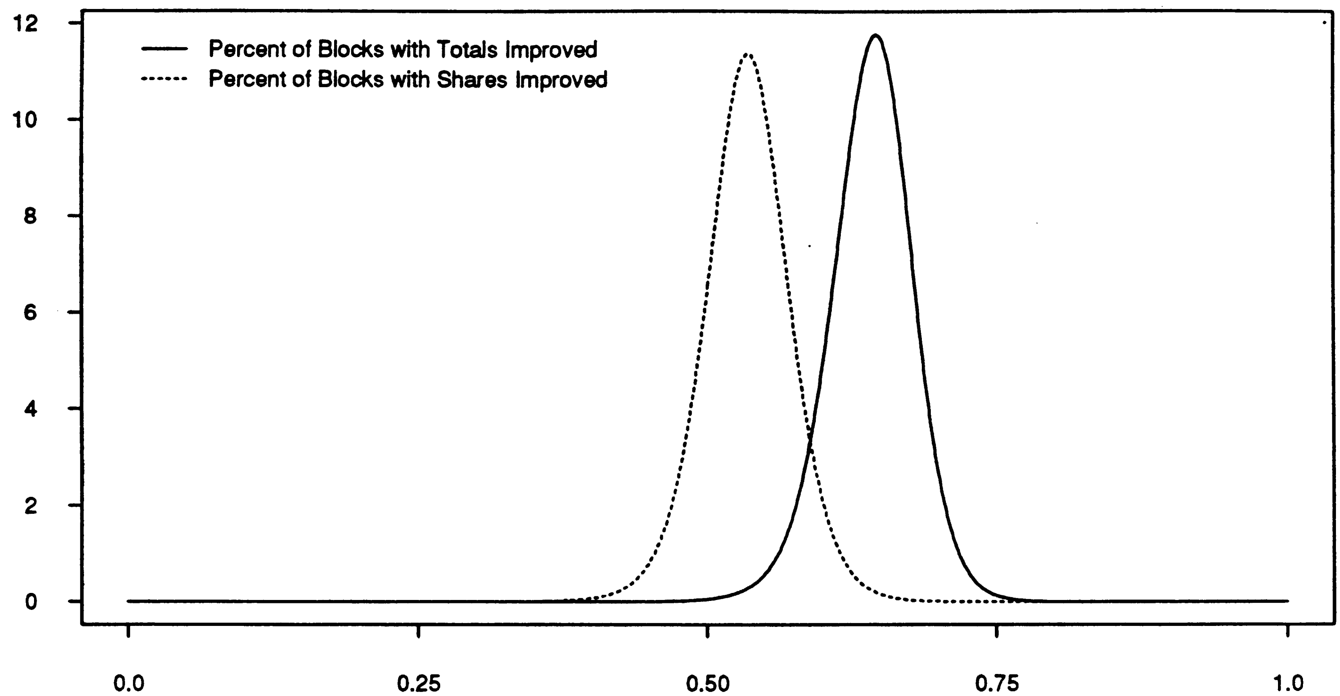
Figure 20: Percent of Blocks Improved, Arlington, With Effects + Int
(250 Simulations)



54% Mean Improvement of Totals, 56% Mean Improvement of Shares

24/250 had < 50% Improved Totals, 16/250 had < 50% Improved Shares

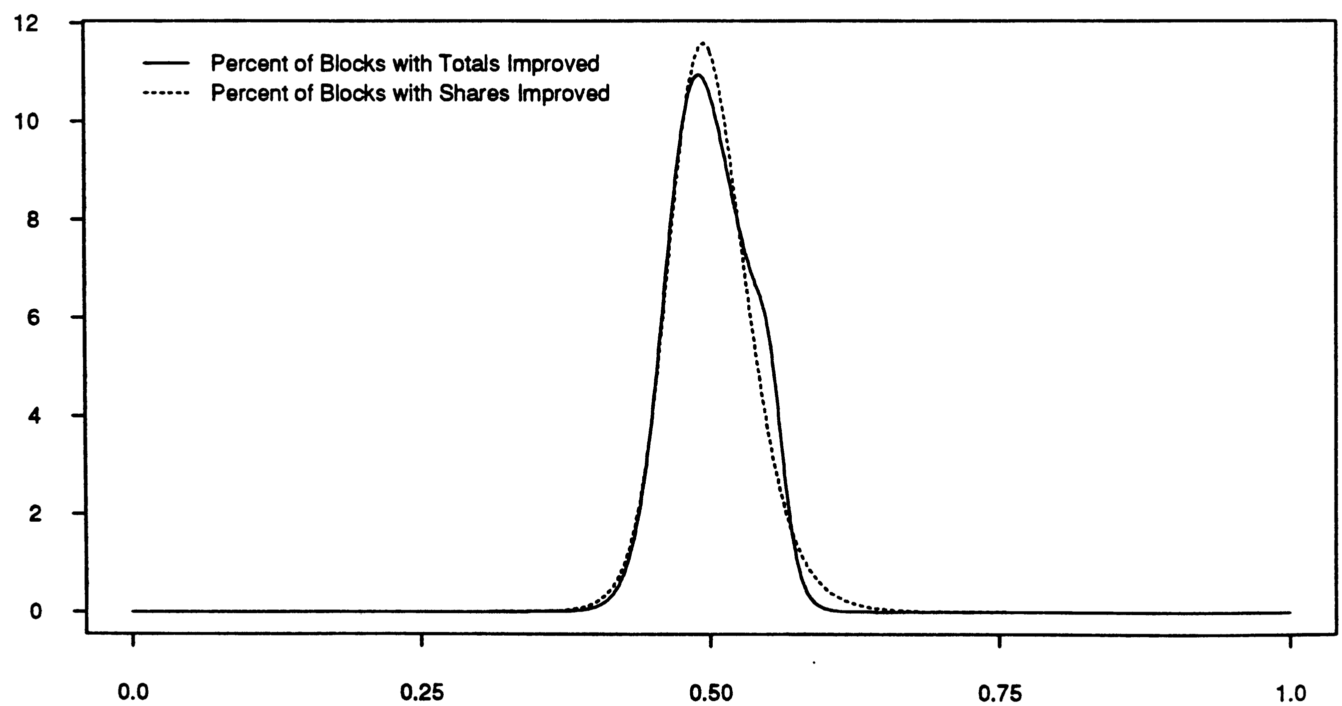
Figure 21: Percent of Blocks Improved, Detago, Without Block Effects
(250 Simulations)



64% Mean Improvement of Totals, 54% Mean Improvement of Shares

0/250 had < 50% Improved Totals, 39/250 had < 50% Improved Shares

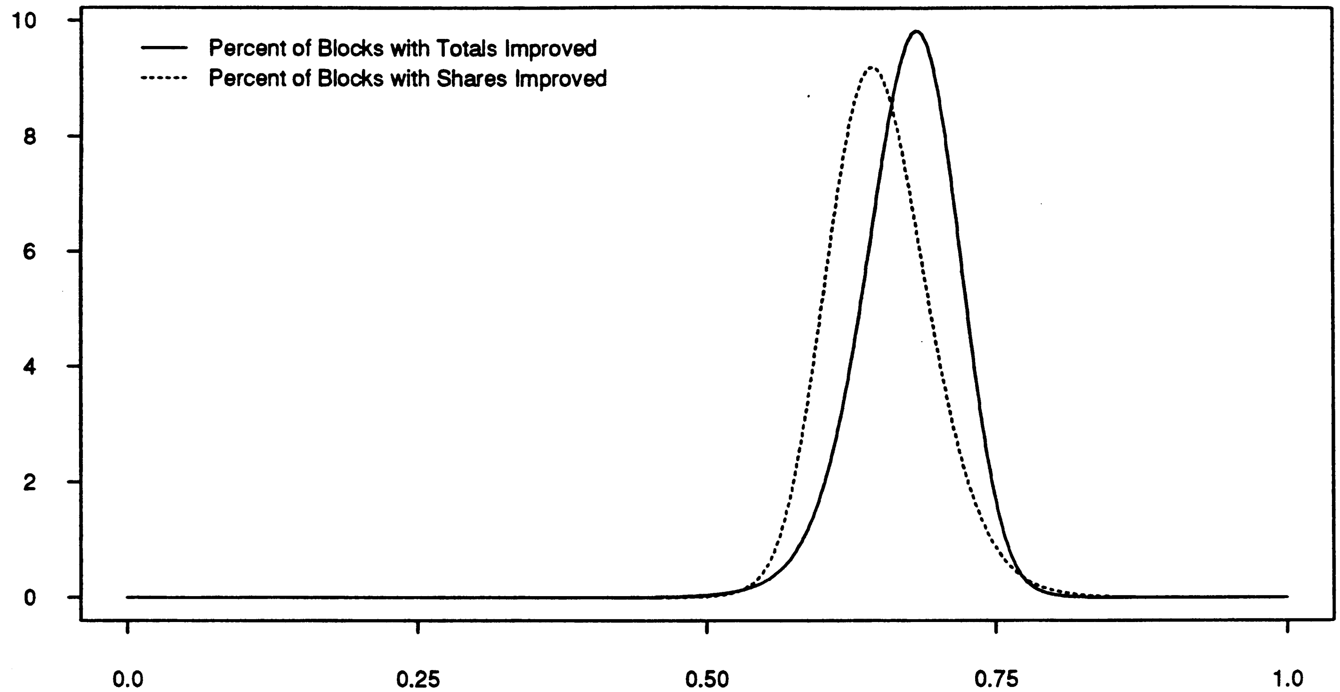
Figure 22: Percent of Blocks Improved, Detago, With Effects + Int
(250 Simulations)



50% Mean Improvement of Totals, 50% Mean Improvement of Shares

114/250 had < 50% Improved Totals, 115/250 had < 50% Improved Shares

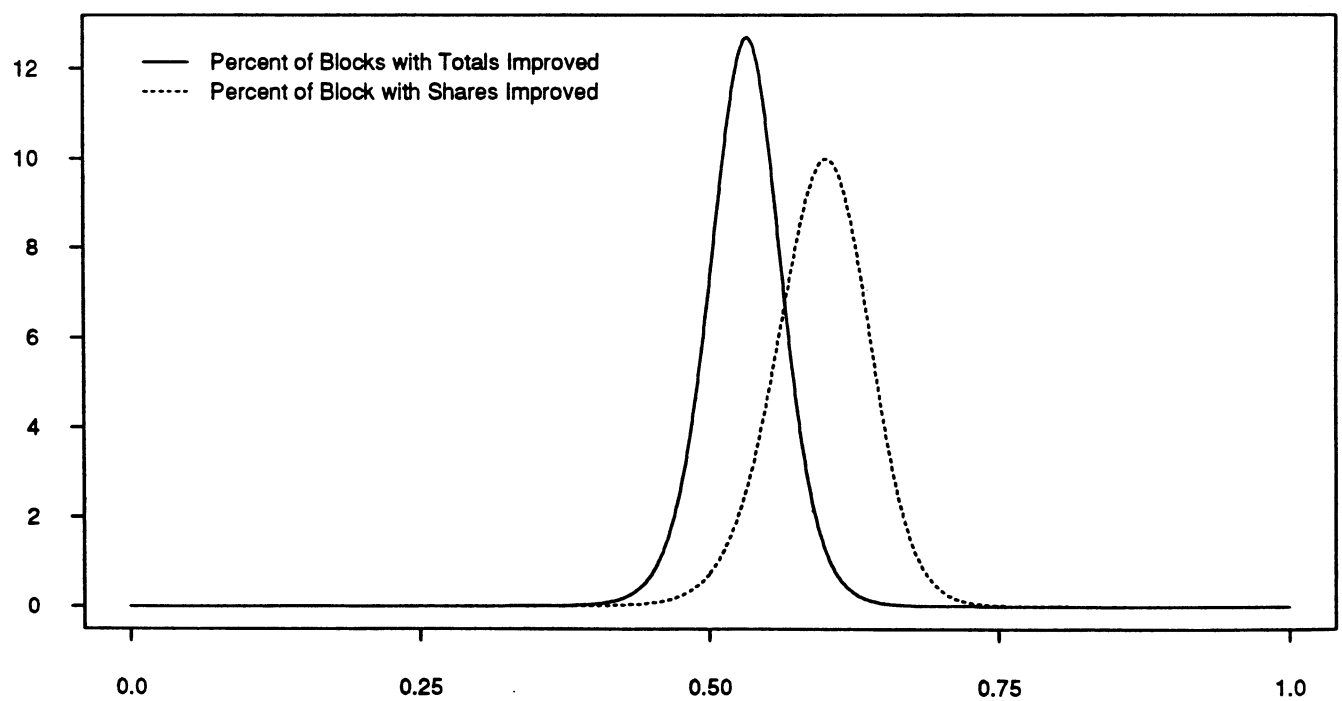
Figure 23: Percent of Blocks Improved, Berkeley, Without Block Effects
(250 Simulations)



67% Mean Improvement of Totals, 65% Mean Improvement of Shares

0/250 had < 50% Improved Totals, 0/250 had < 50% Improved Shares

Figure 24: Percent of Blocks Improved, Berkeley, With Effects + Int



53% Mean Improvement of Totals, 60% Mean Improvement of Shares

45/250 had < 50% Improved Totals, 4/250 had < 50% Improved Shares

Totals

	Districts	Blocks
New York	63.0%	50.0%
Arlington	62.0%	54.0%
Detago	56.0%	50.0%
Berkeley	62.0%	53.0%

Shares

	Districts	Blocks
New York	63.0%	51.0%
Arlington	53.0%	56.0%
Detago	52.0%	50.0%
Berkeley	67.0%	60.0%

Tables 7 and 8: Mean Proportion of Districts and Blocks Improved under FIT3 (250 Simulations)

References

1. Kooperberg, C. and Stone, C. J. (1991). A study of logspline density estimation. *Comp. Statist. Data Anal.* **12** 327-347.

2. Kooperberg, C. and Stone, C. J. (1992). Logspline density estimation for censored data. Technical Report No. 226, University of Washington Department of Statistics.

VI. Appendix: Miscellaneous Figures

Figure A1: Predicted vs. Actual GO's, Detago,
Fits with and without district effects.

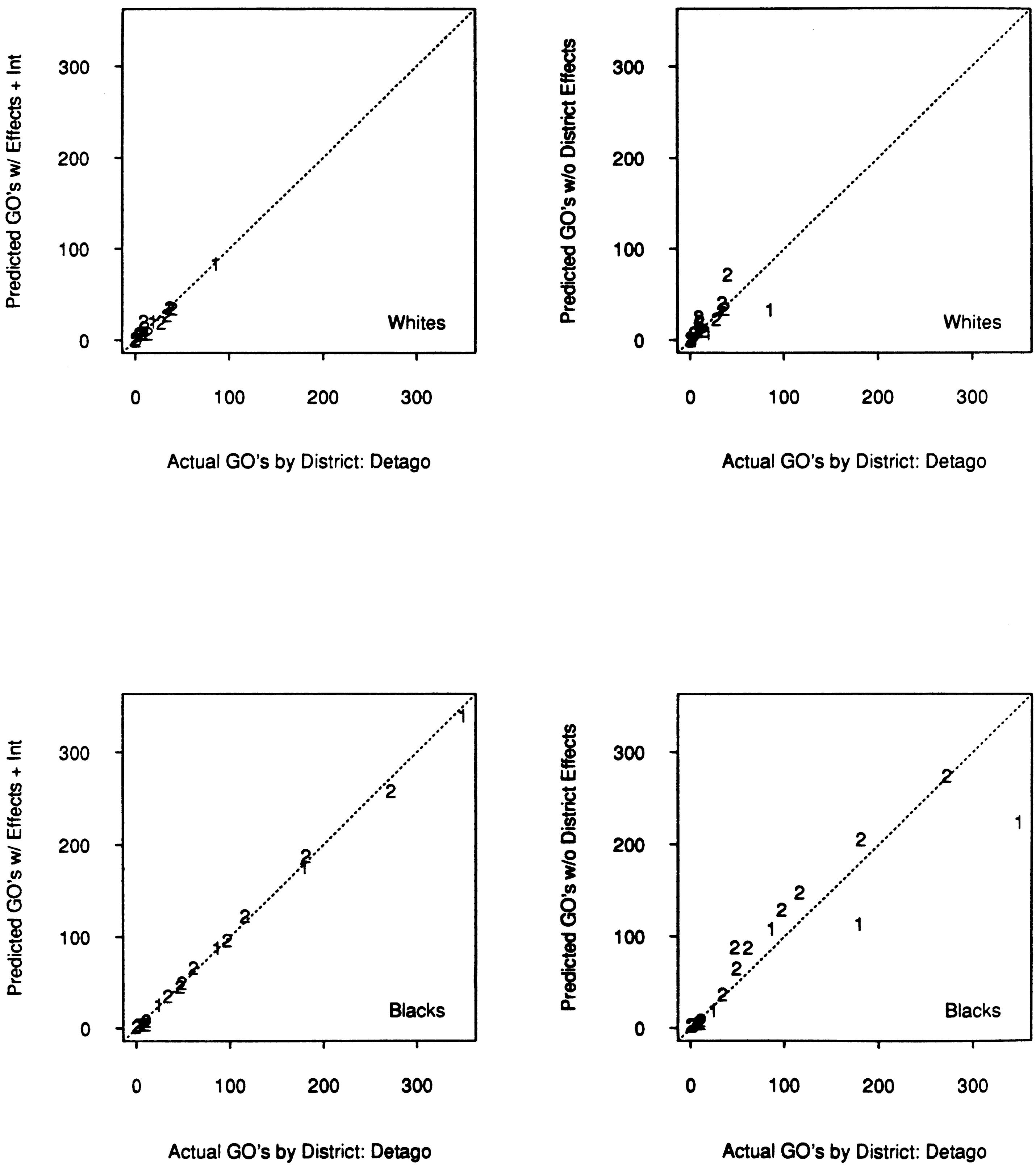


Figure A1 (cont): Predicted vs. Actual GO's, Detago,
Fits with and without district effects.

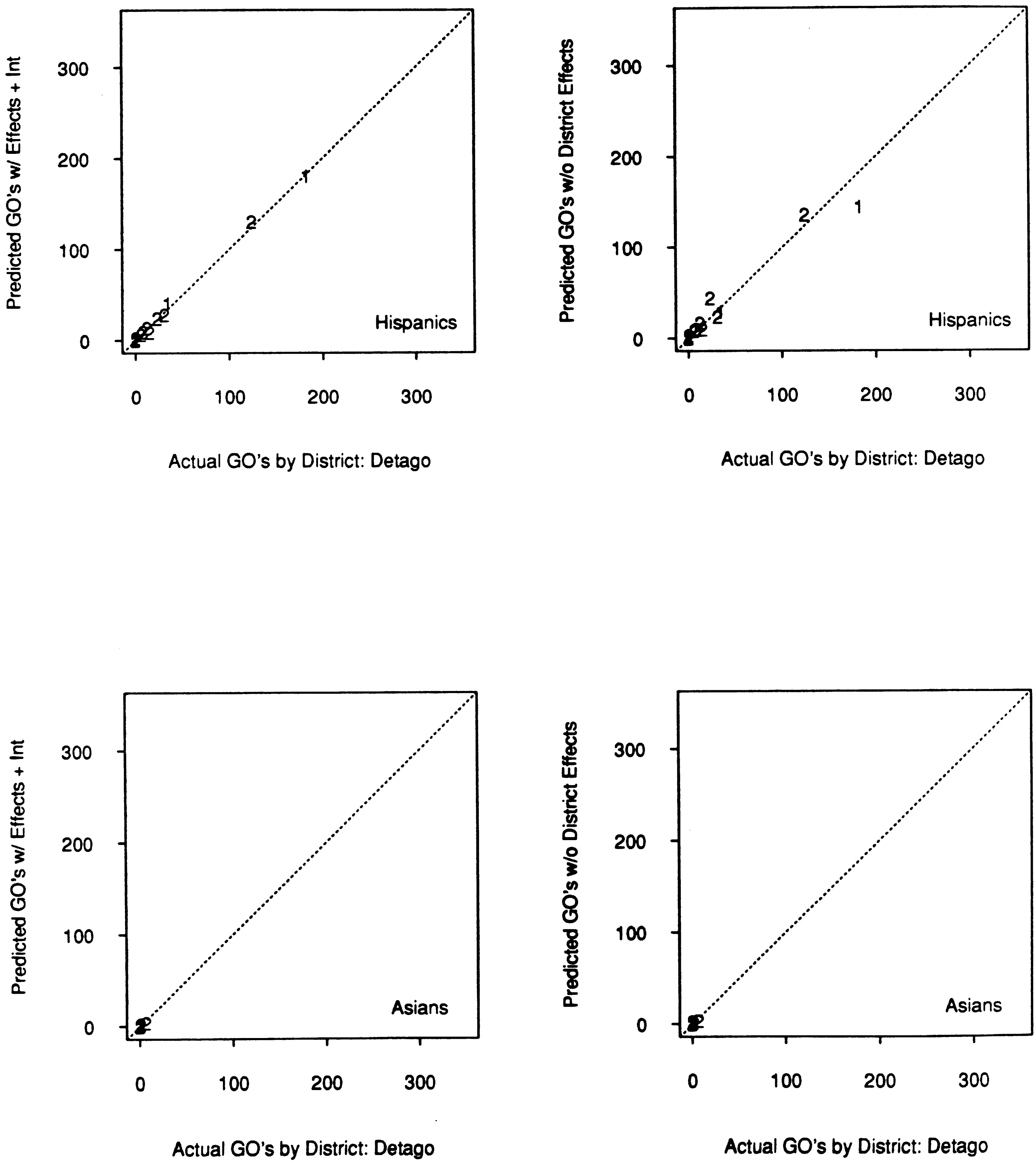


Figure A2: Predicted vs. Actual EE's, Detago,
Fits with and without district effects.

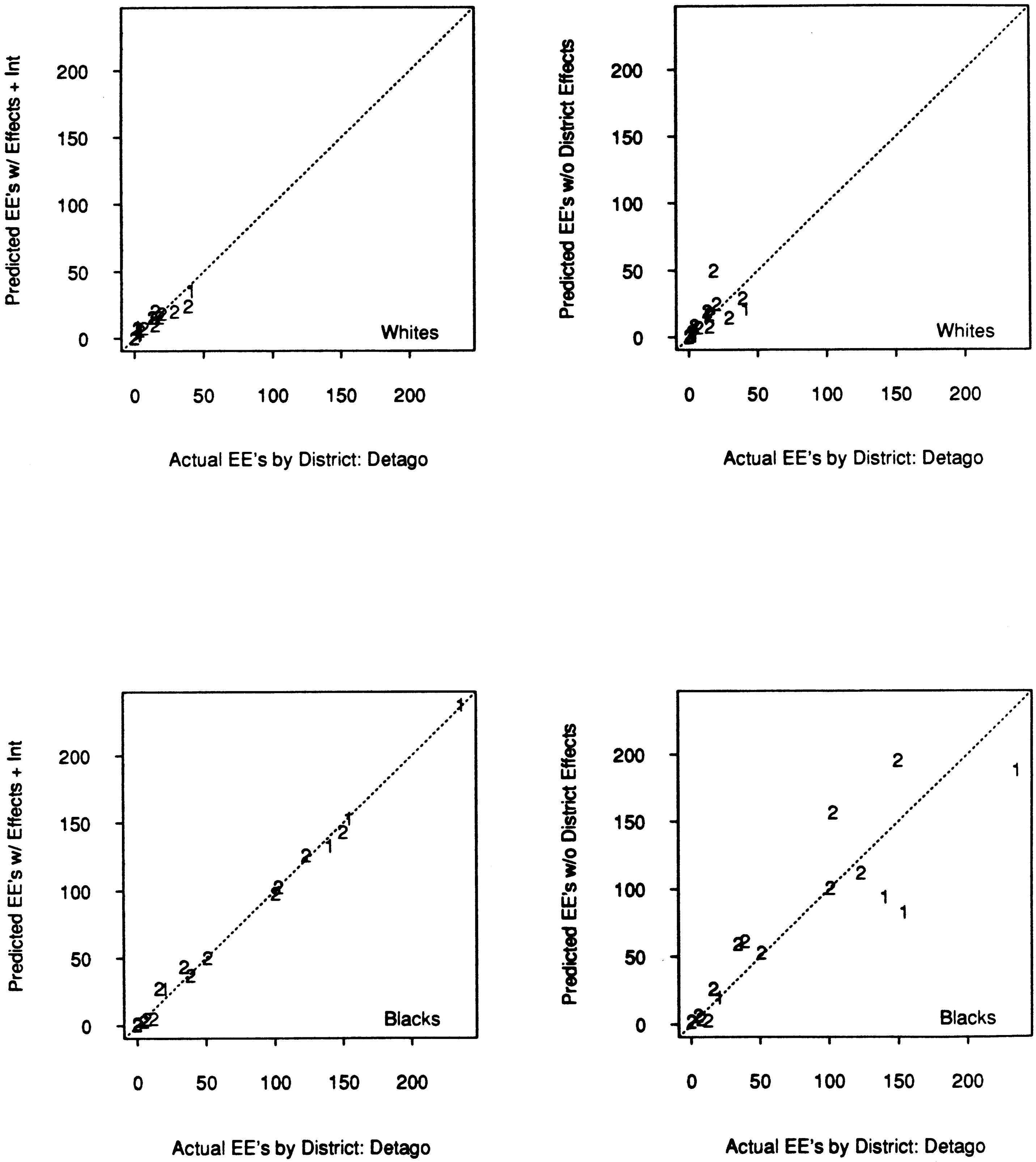


Figure A2 (cont): Predicted vs. Actual EE's, Detago,
Fits with and without district effects.

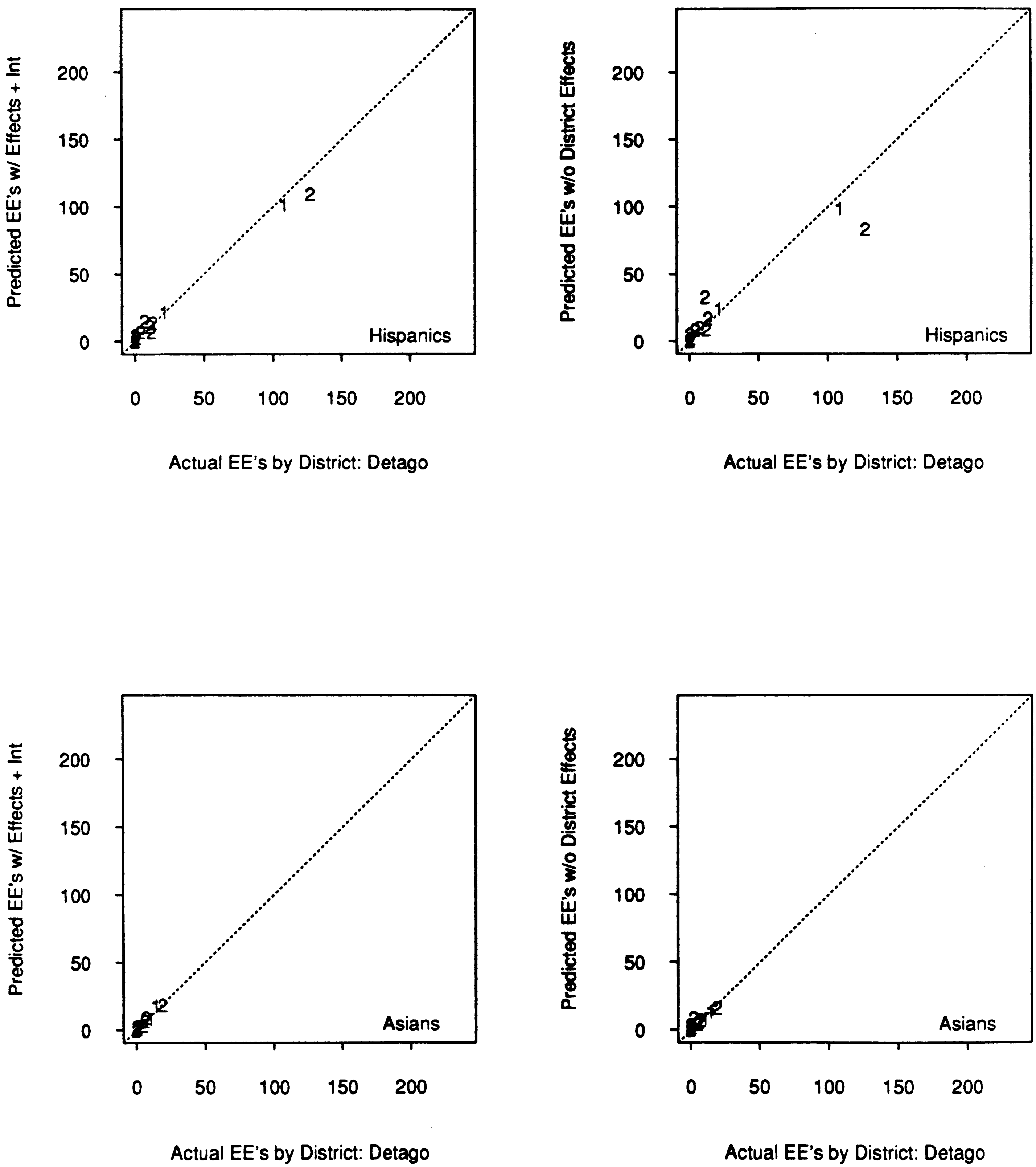


Figure A3: Predicted vs. Actual GO's, New York,
Fits with and without district effects.

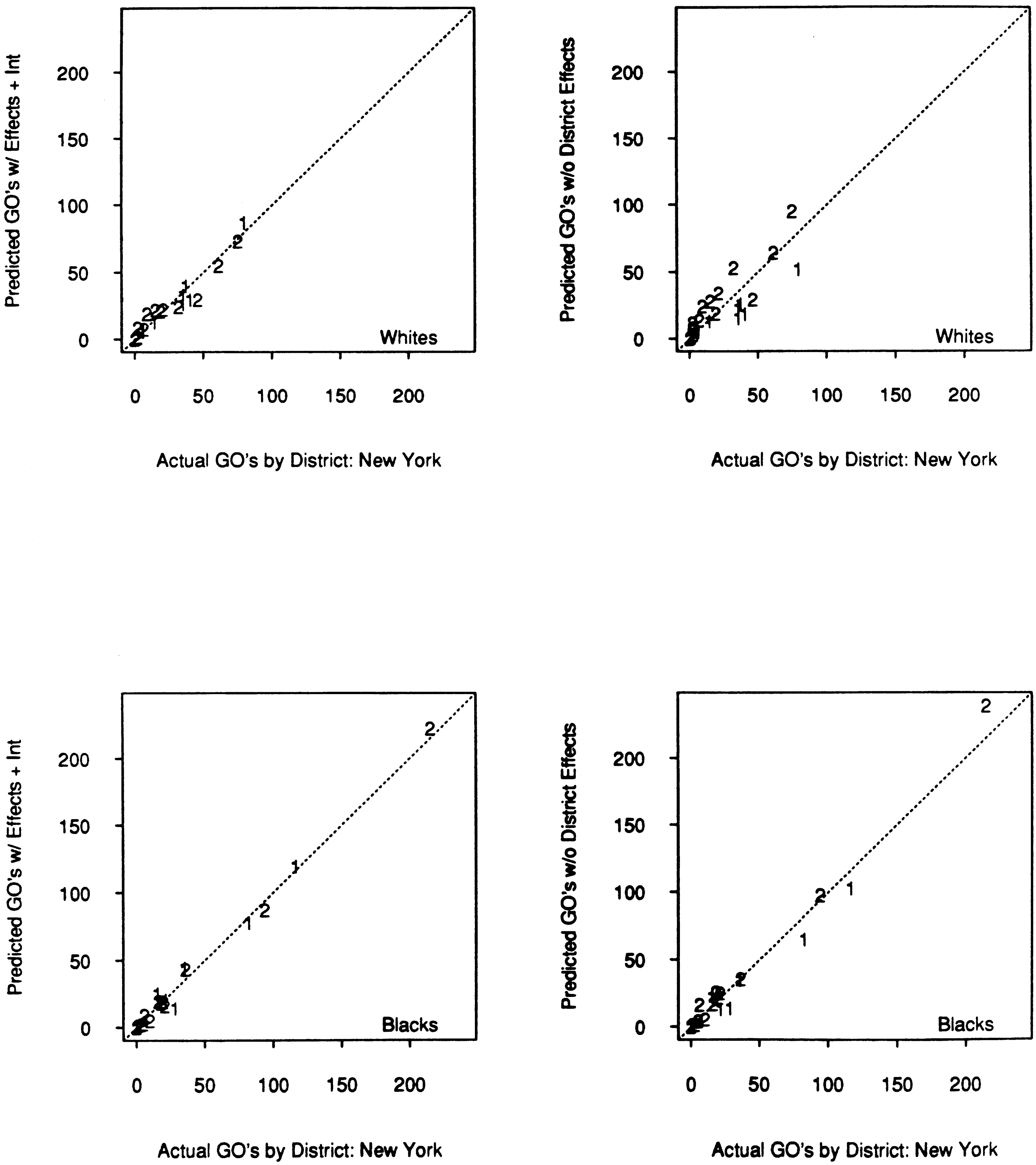


Figure A3 (cont): Predicted vs. Actual GO's, New York,
Fits with and without district effects.

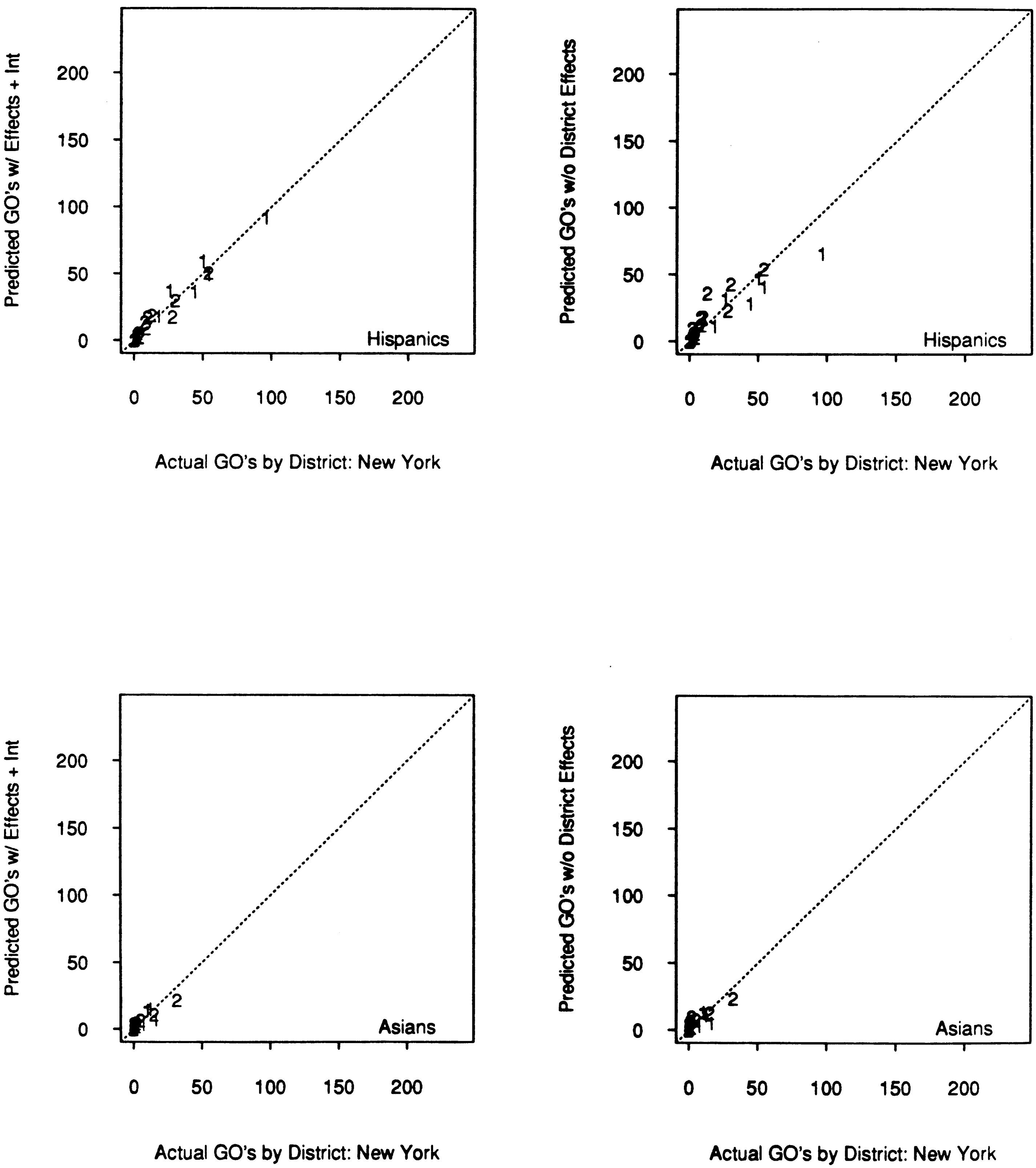


Figure A4: Predicted vs. Actual EE's, New York,
Fits with and without district effects.

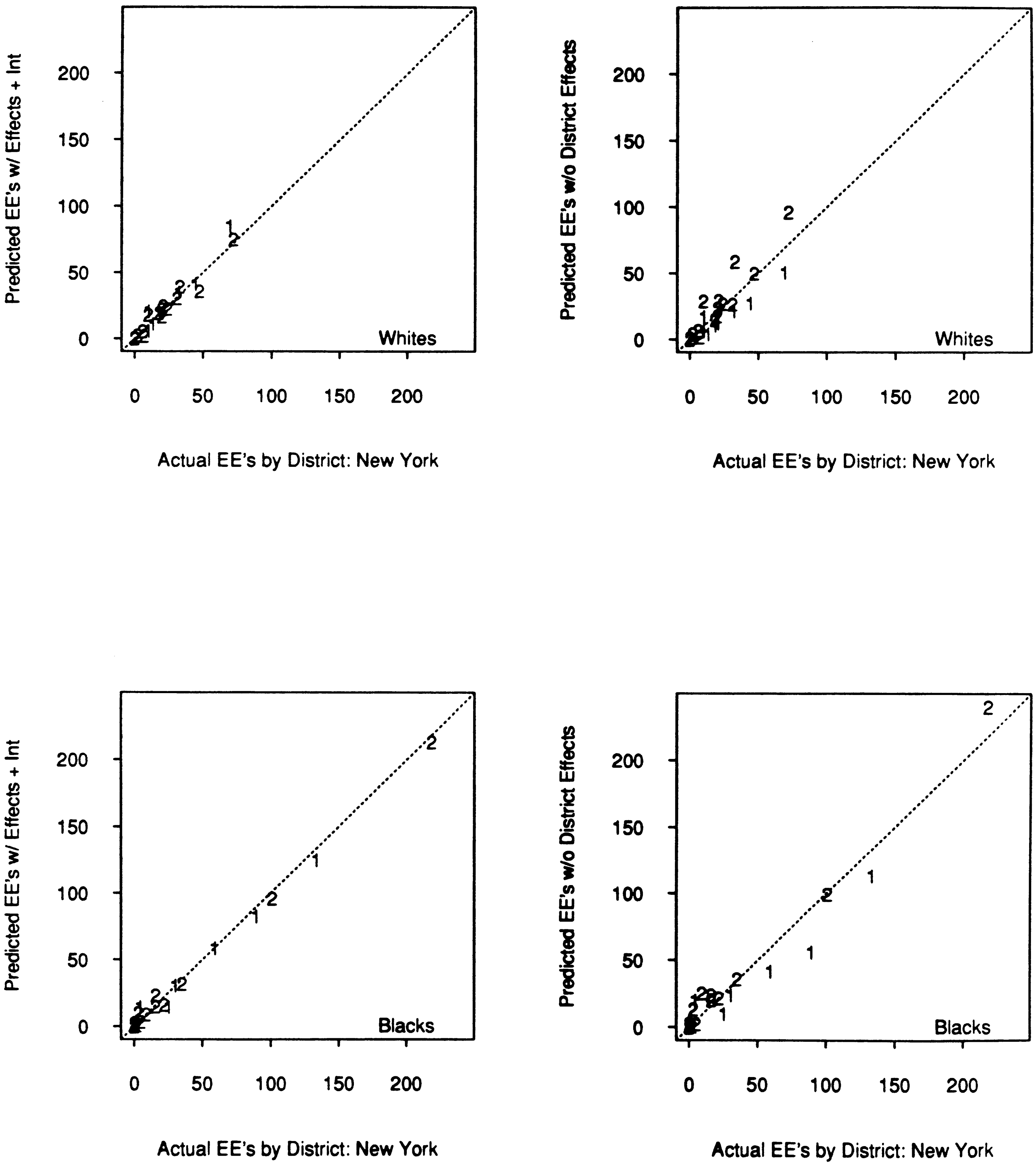


Figure A4 (cont): Predicted vs. Actual EE's, New York,
Fits with and without district effects.

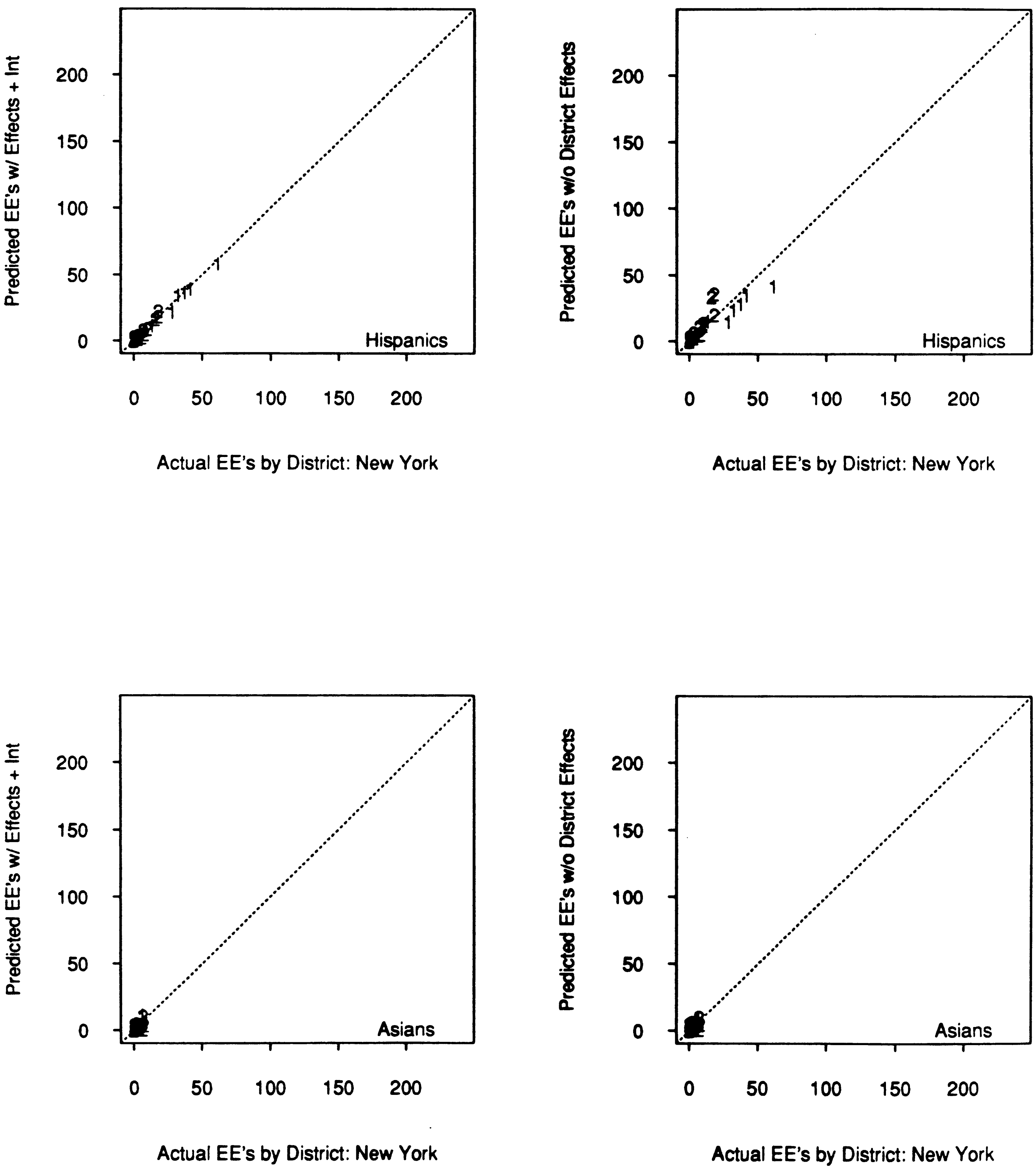
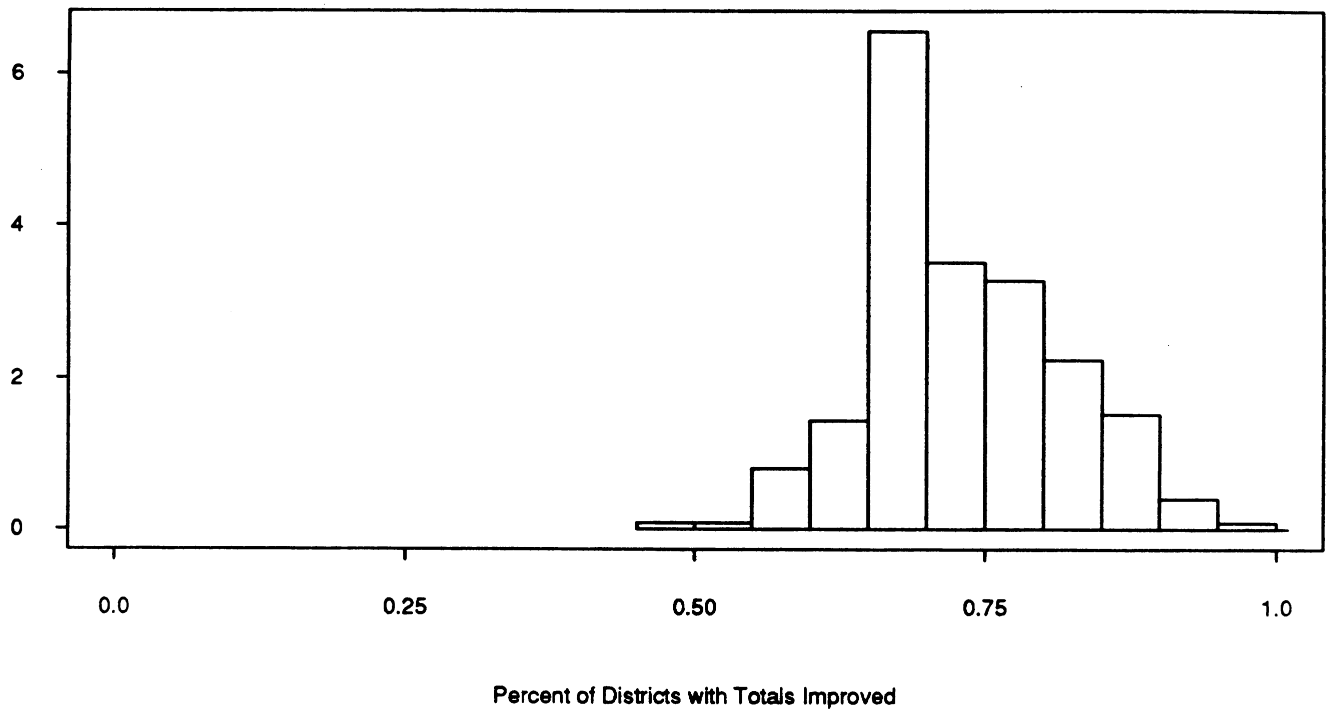


Figure A5: Percent of Districts Improved, New York, Without District Effects
(250 Simulations)



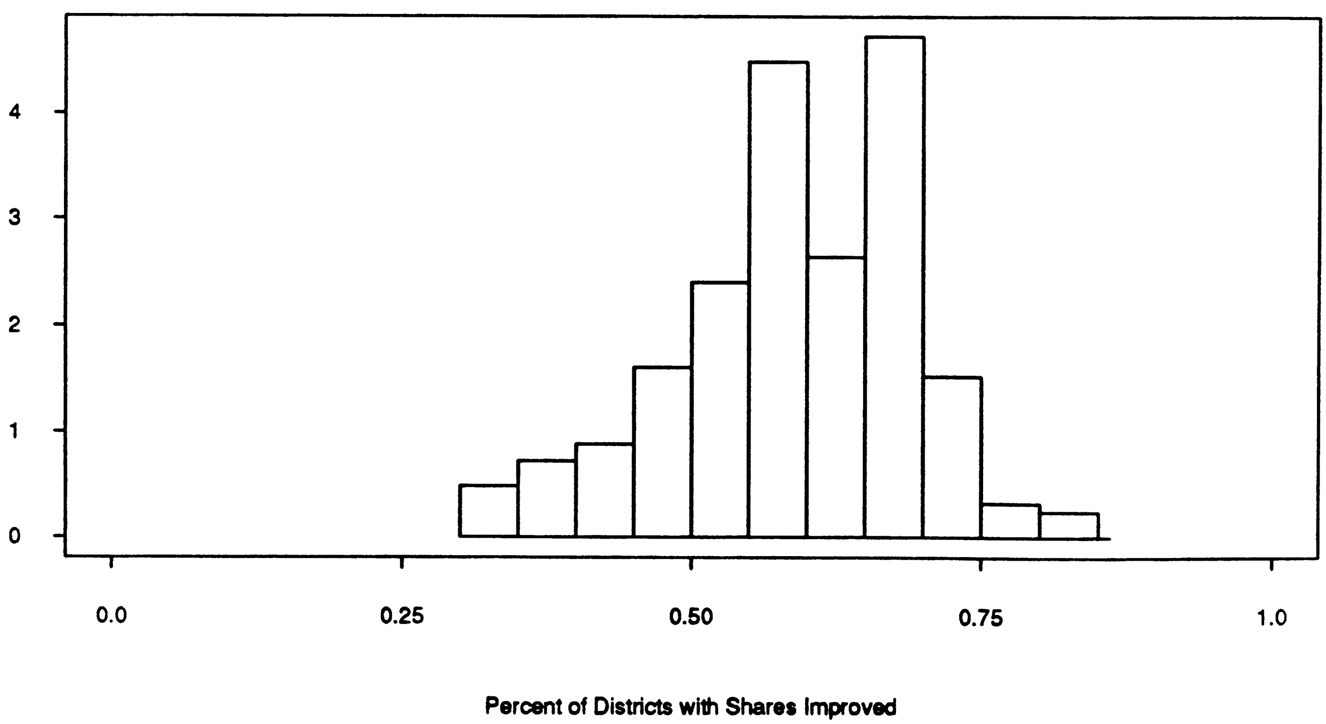
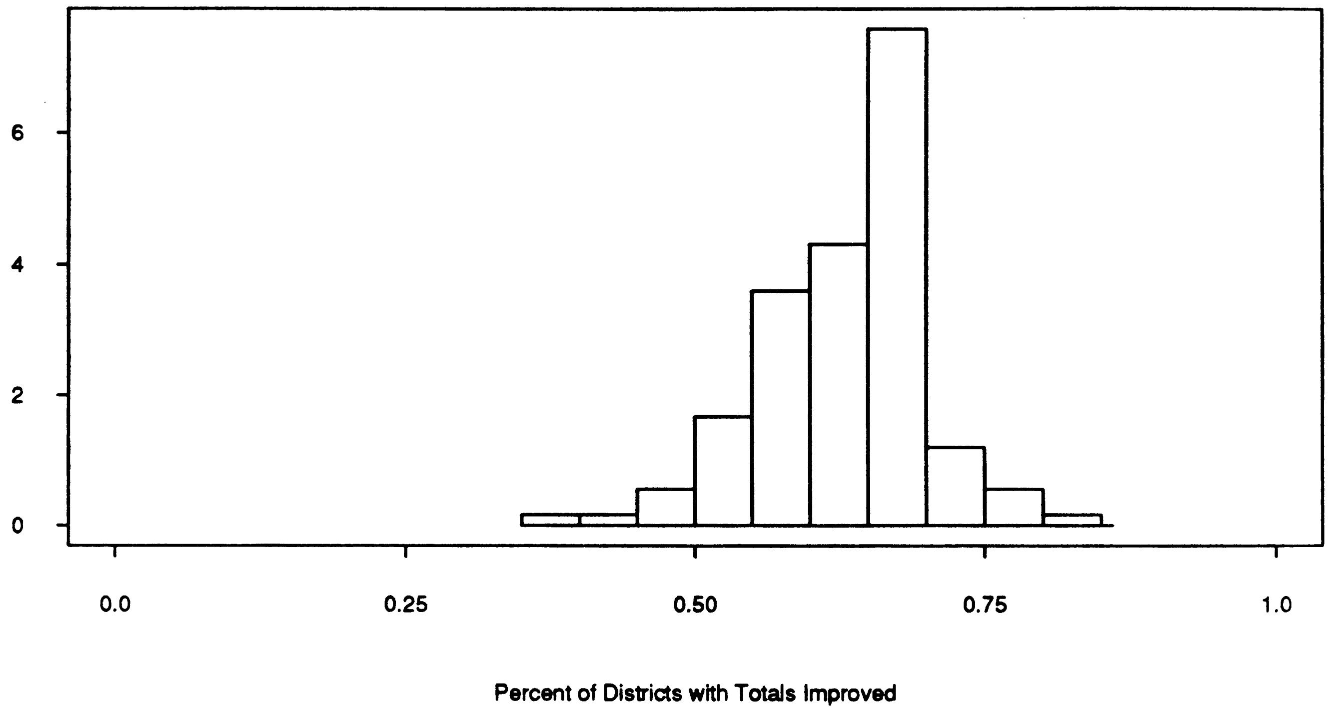


Figure A6: Percent of Districts Improved, New York, With Effects + Int
(250 Simulations)



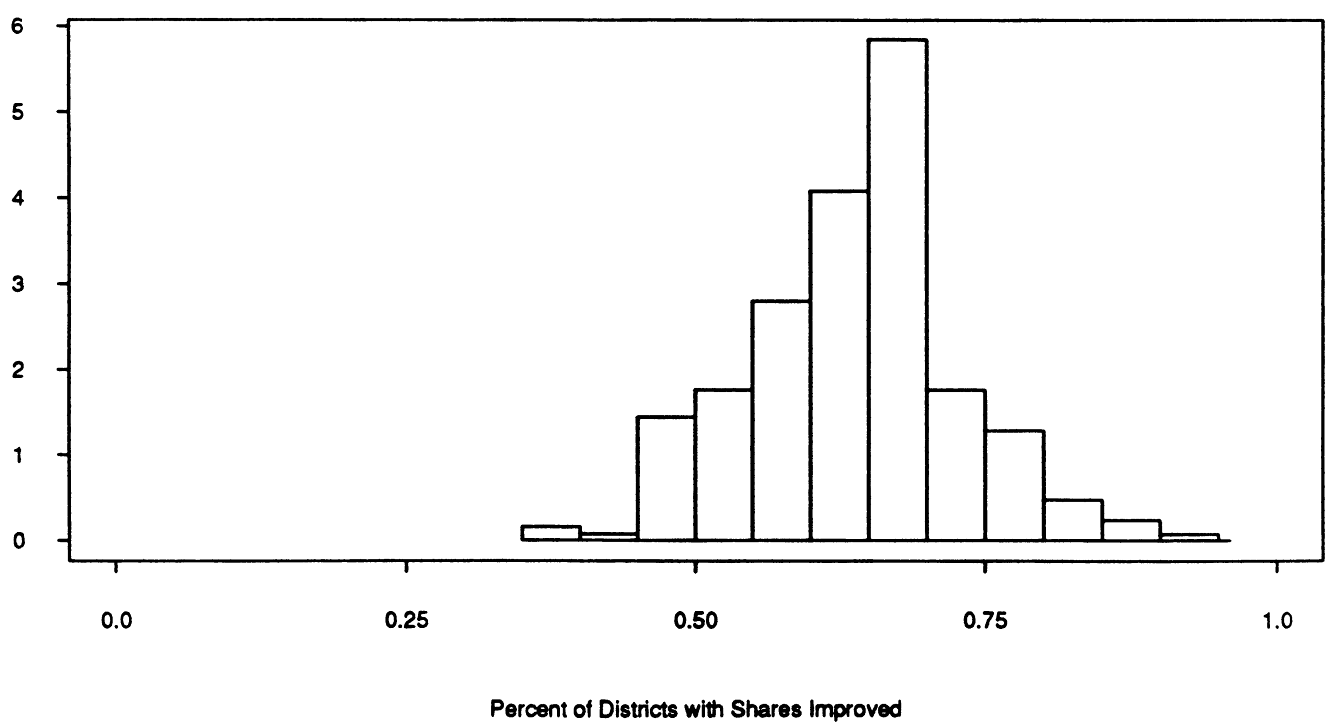
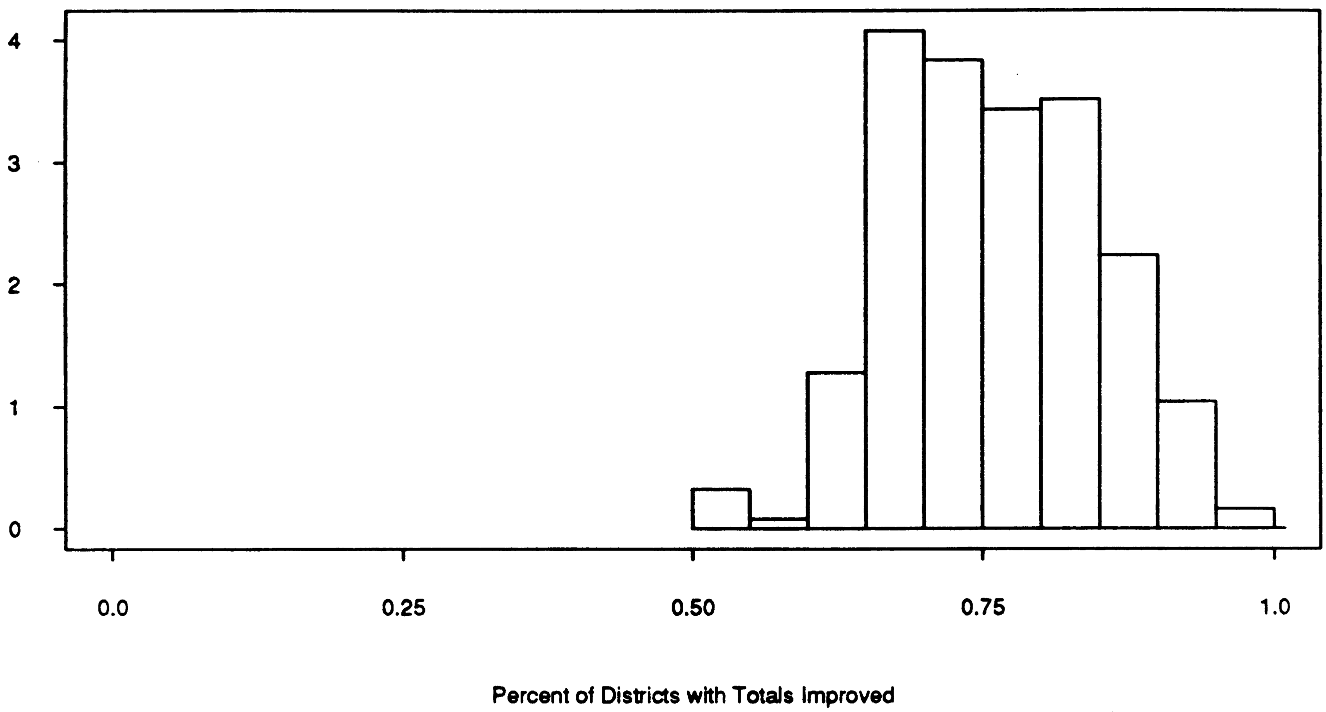


Figure A7: Percent of Districts Improved, Arlington, Without District Effects
(250 Simulations)



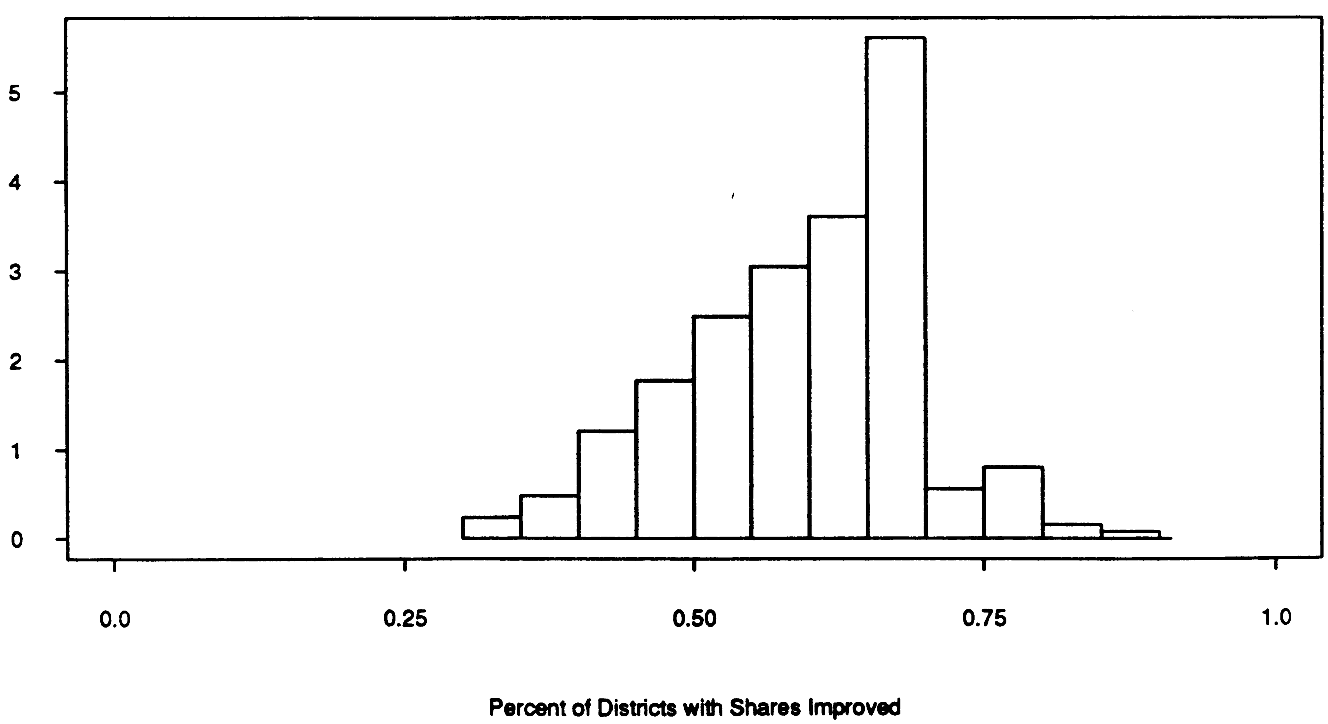
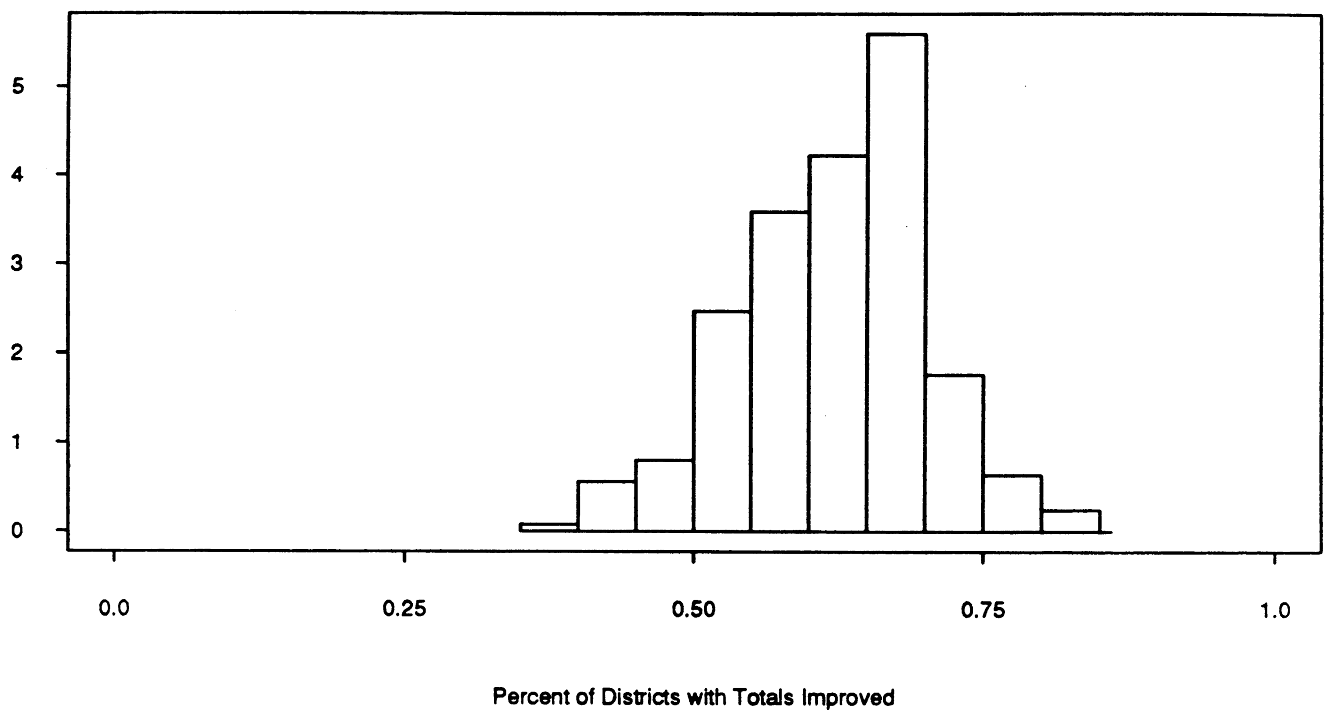


Figure A8: Percent of Districts Improved, Arlington, With Effects + Int
(250 Simulations)



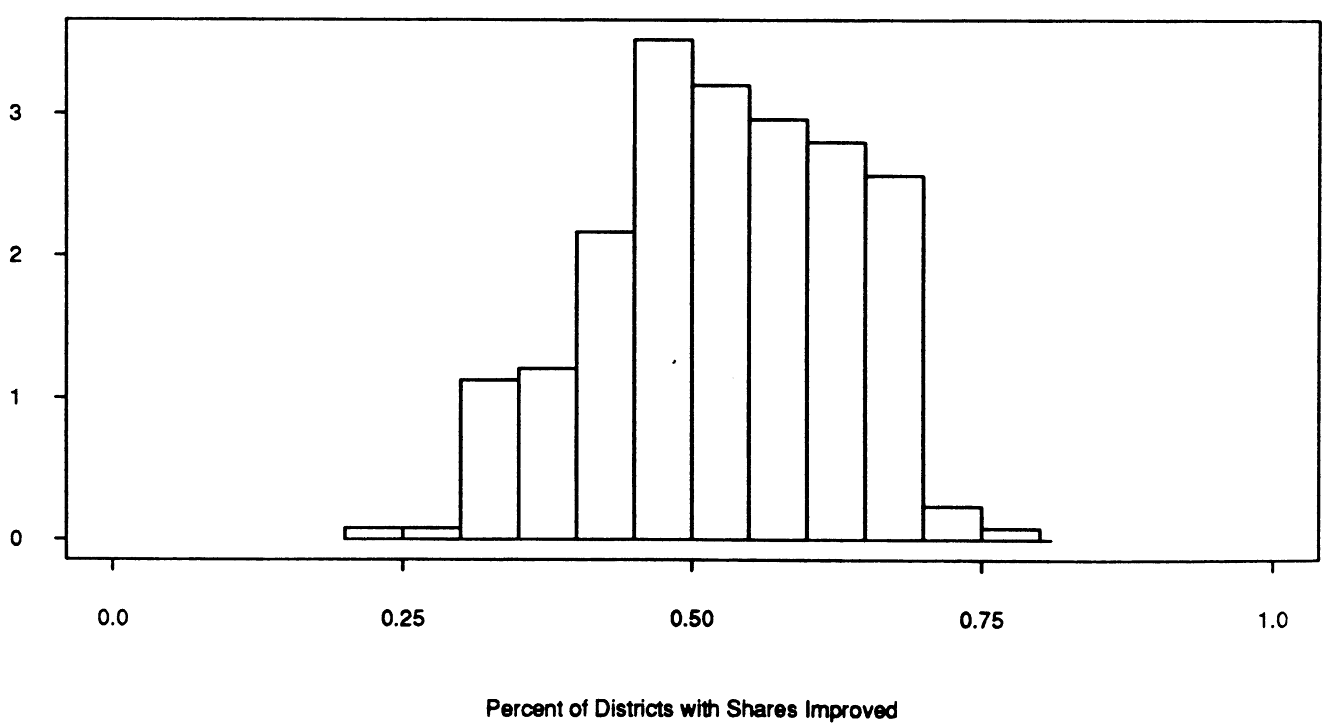
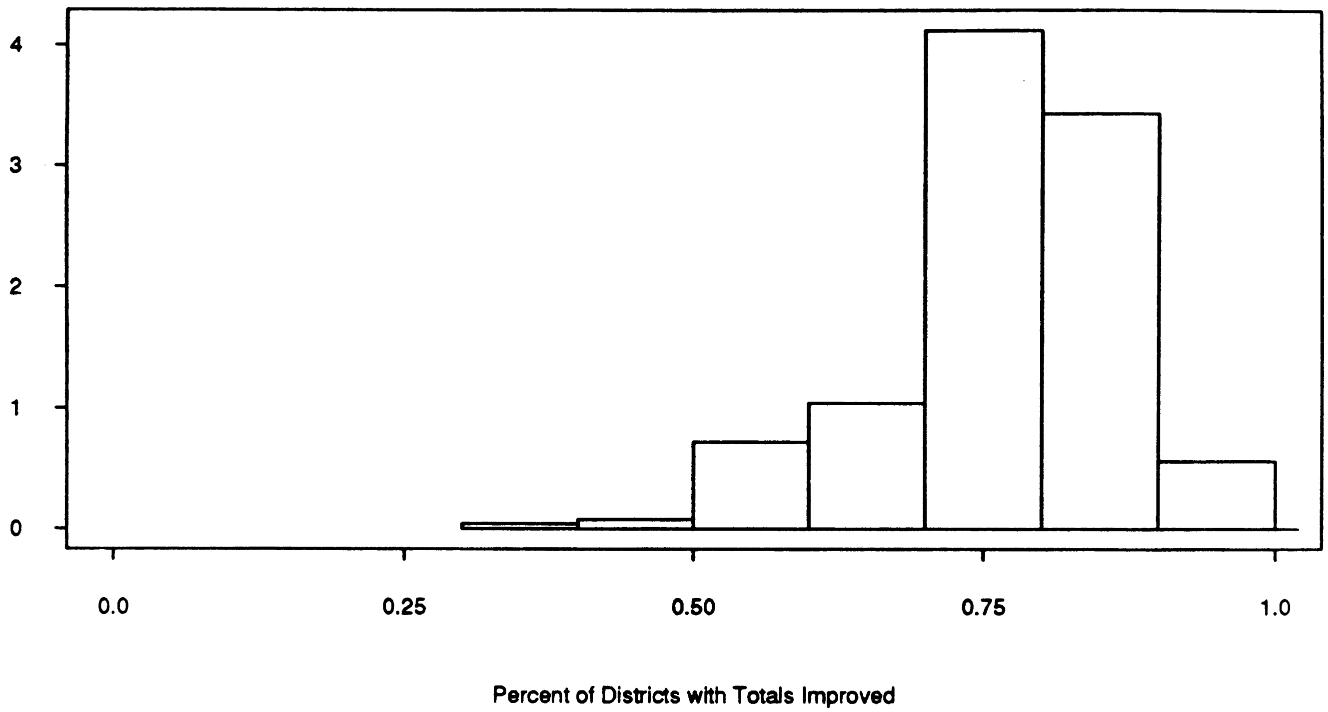


Figure A9: Percent of Districts Improved, Detago, Without District Effects
(250 Simulations)



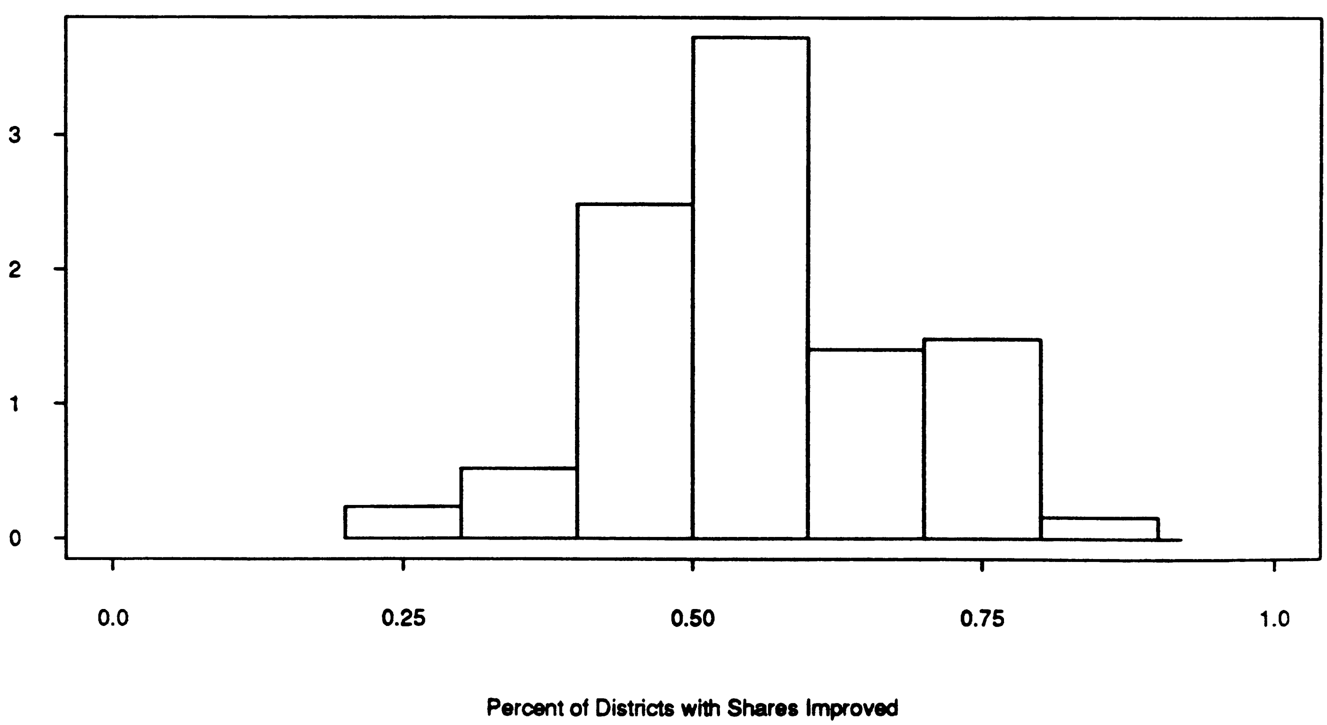
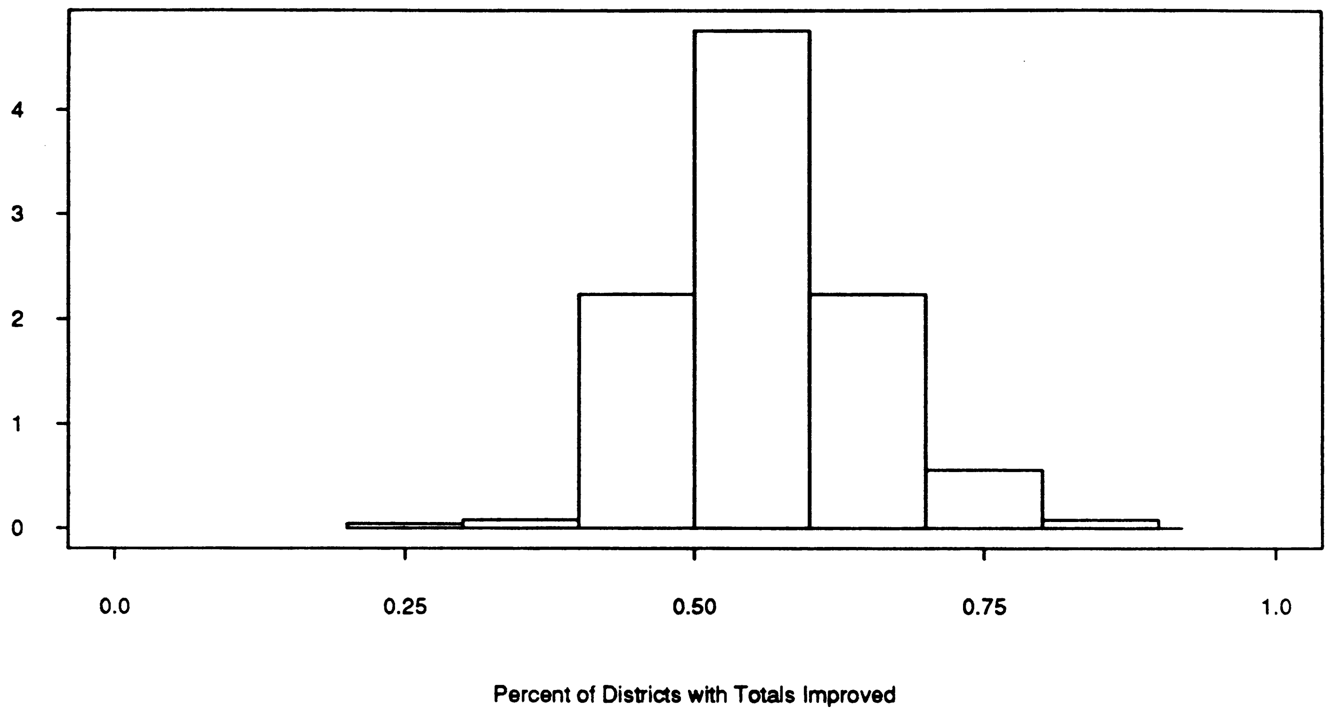


Figure A10: Percent of Districts Improved, Detago, With Effects + Int
(250 Simulations)



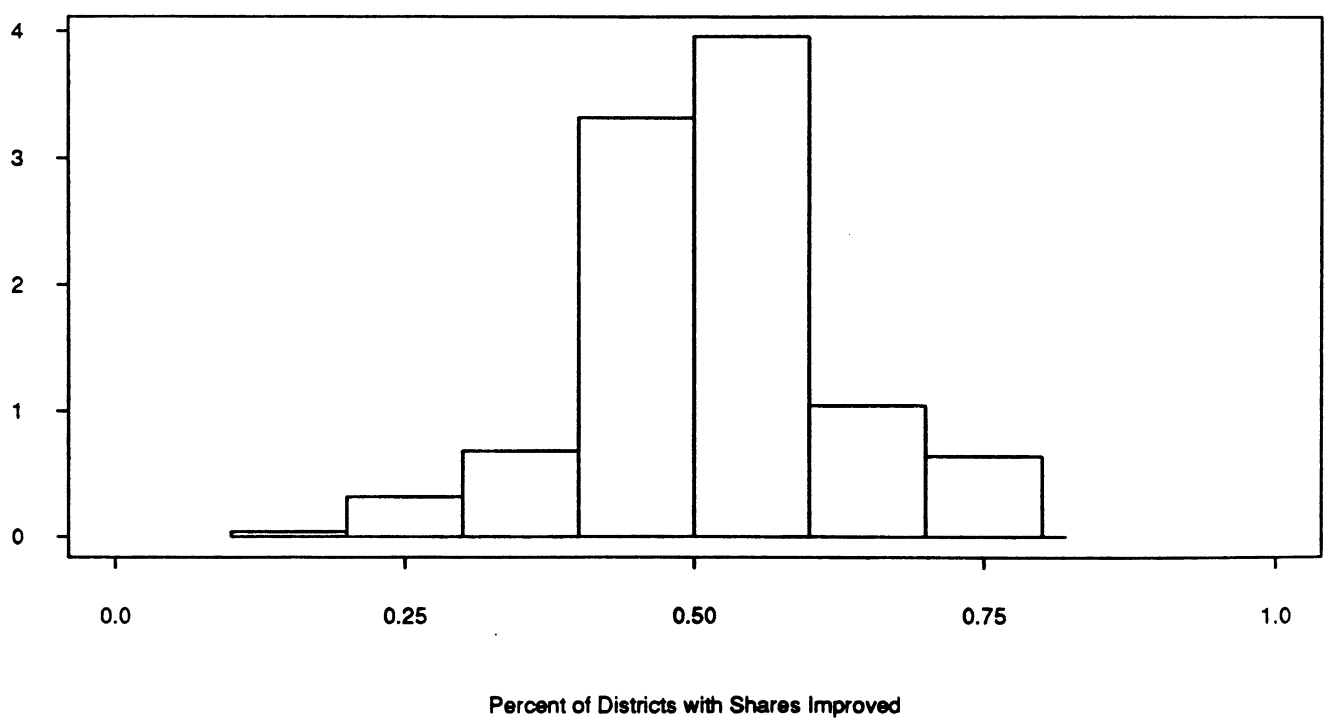
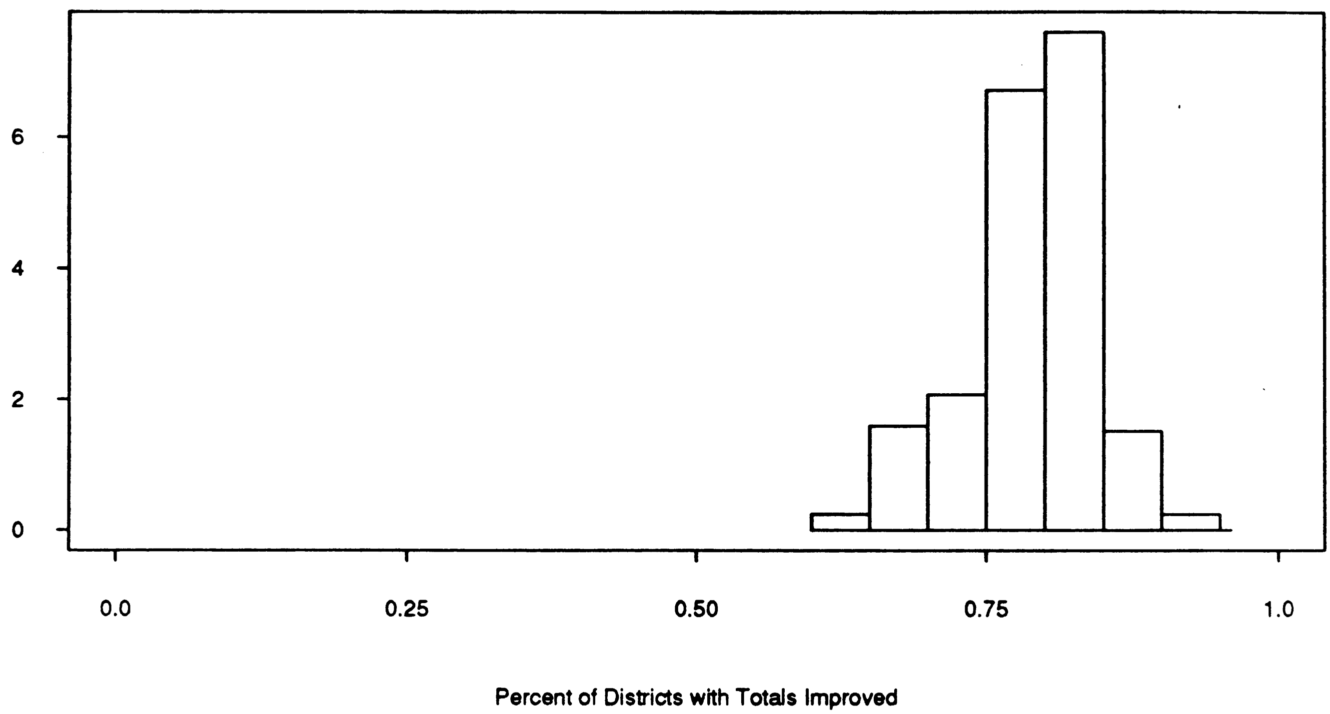


Figure A11: Percent of Districts Improved, Berkeley, Without District Effects
(250 Simulations)



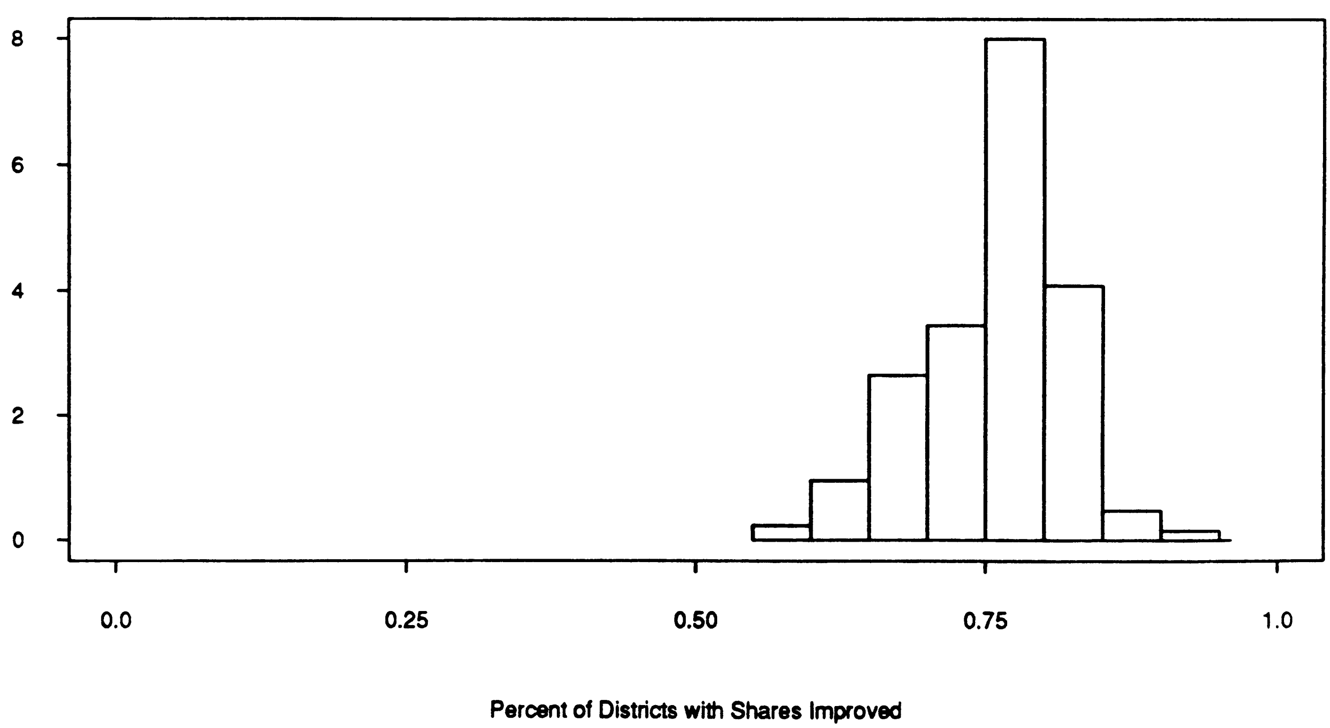
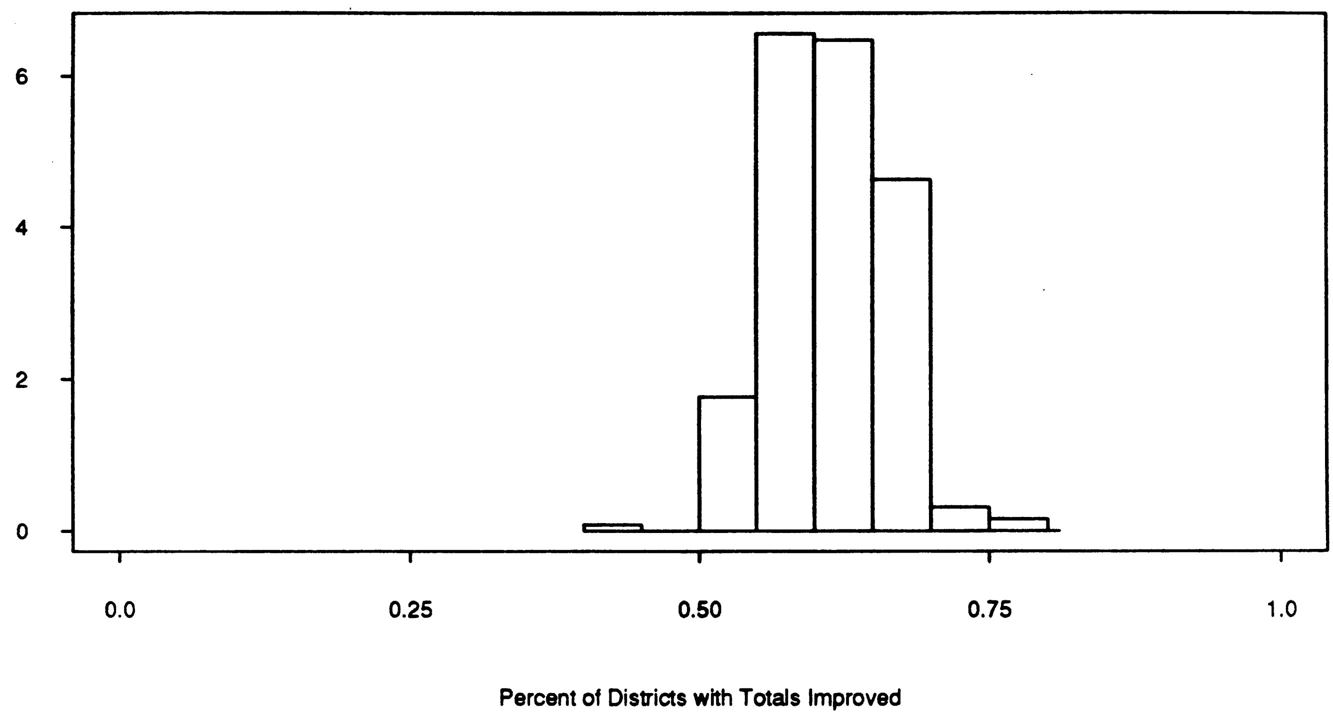


Figure A12: Percent of Districts Improved, Berkeley, With Effects + Int
(250 Simulations)



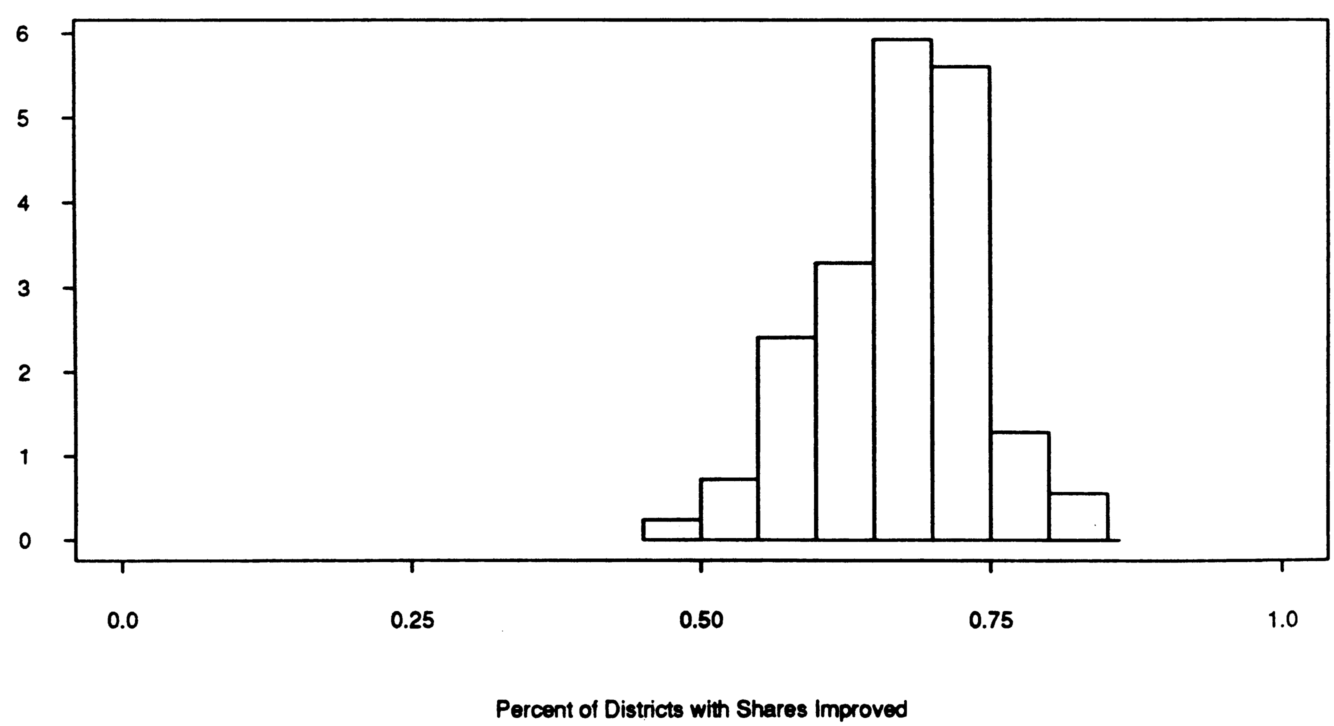
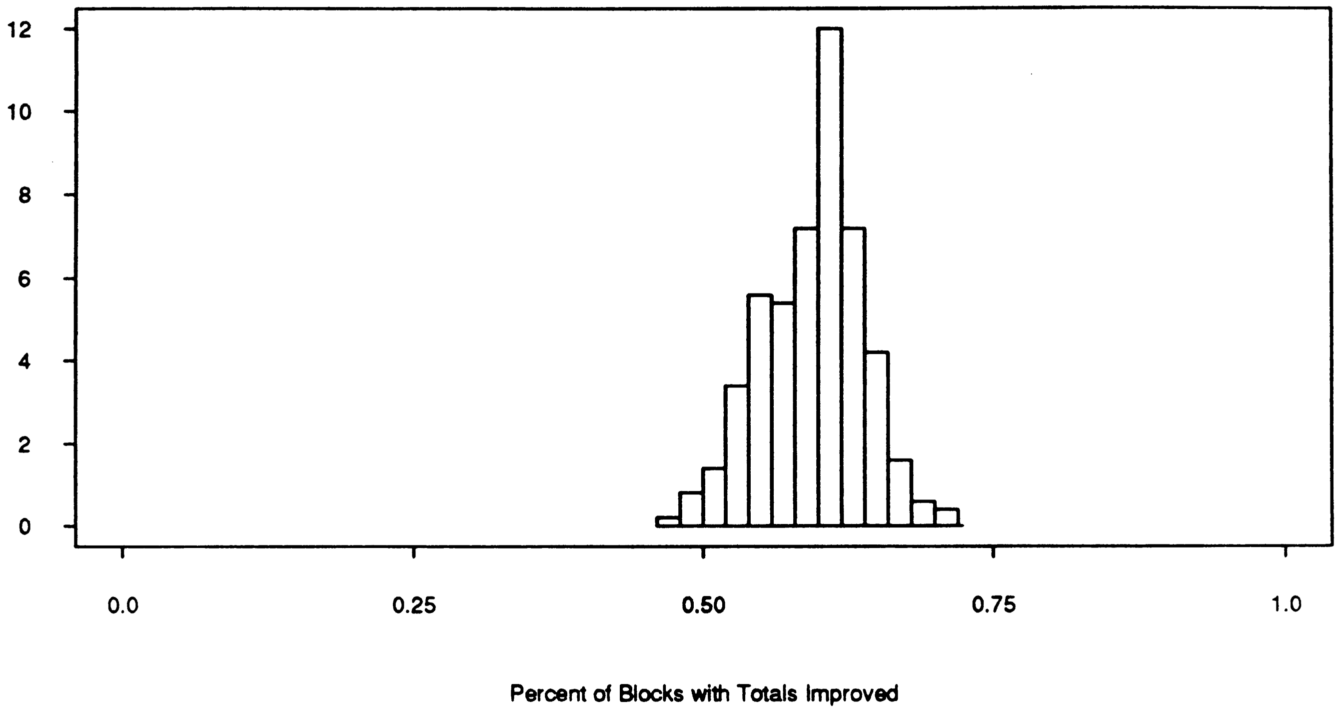


Figure A13: Percent of Blocks Improved, New York, Without Block Effects
(250 Simulations)



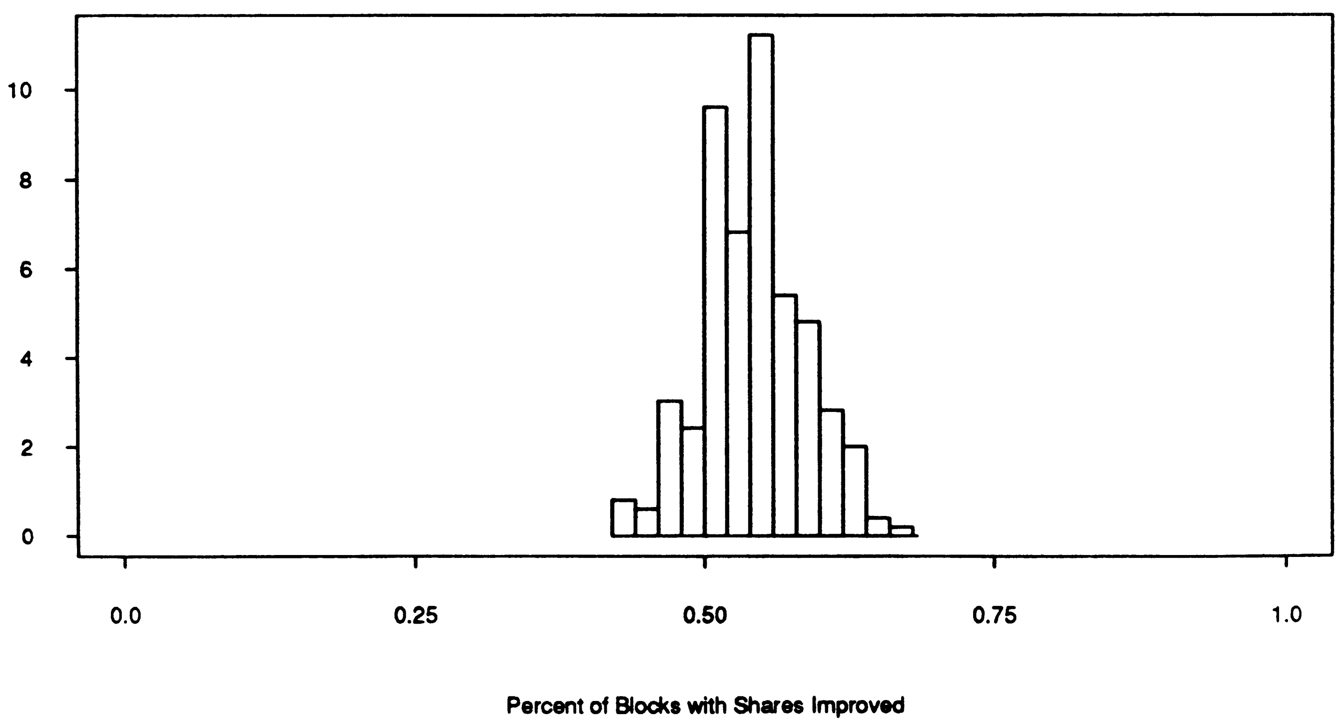
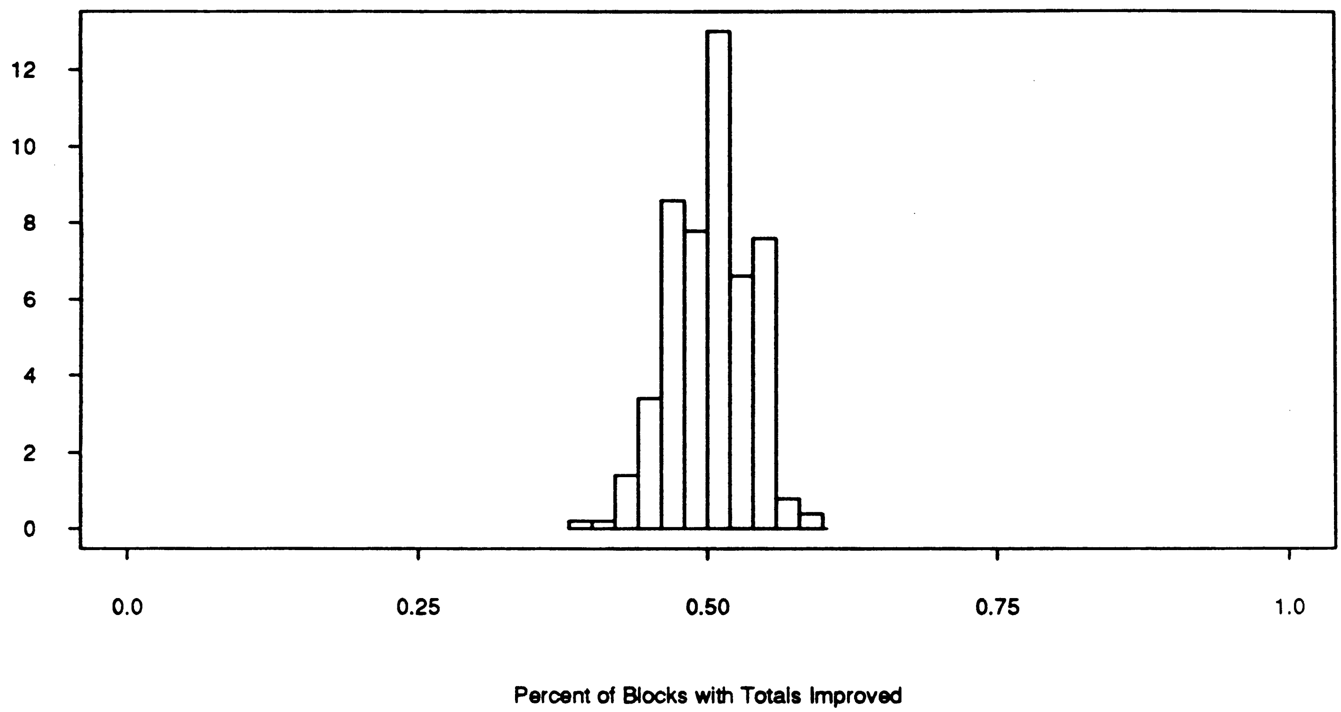


Figure A14: Percent of Blocks Improved, New York, With Effects + Int
(250 Simulations)



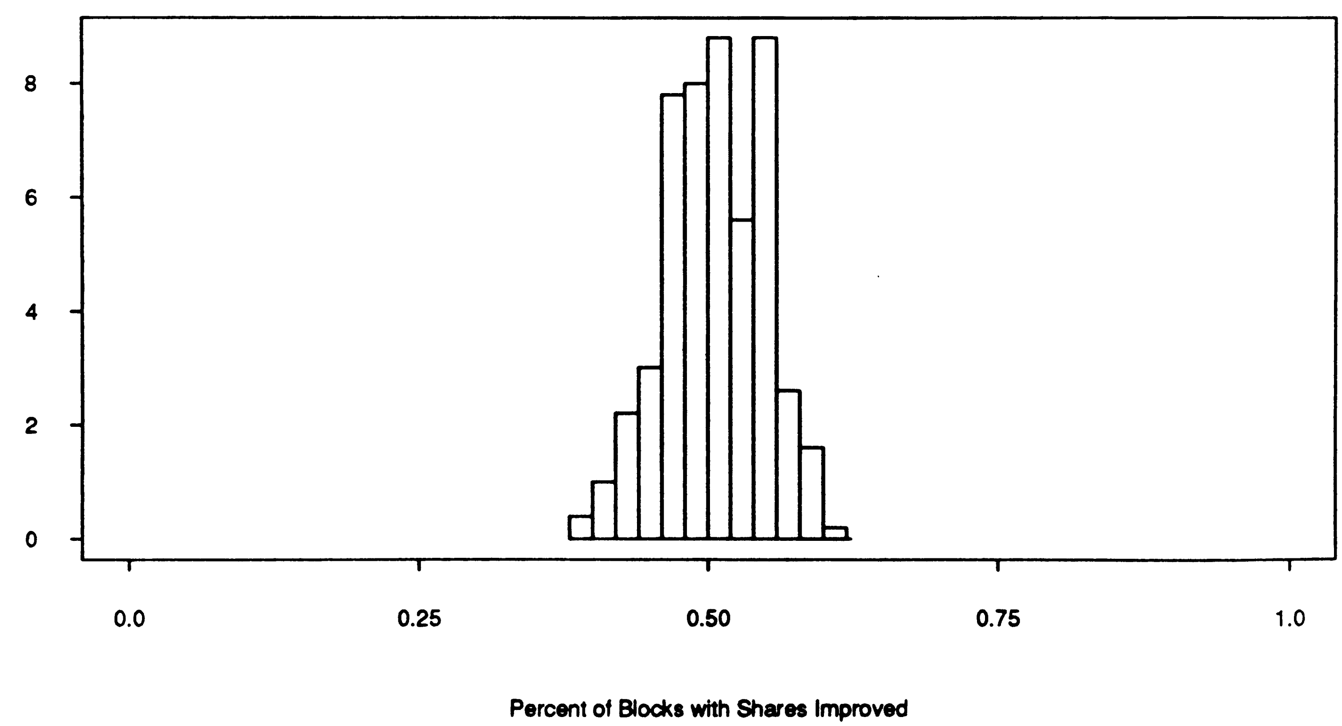
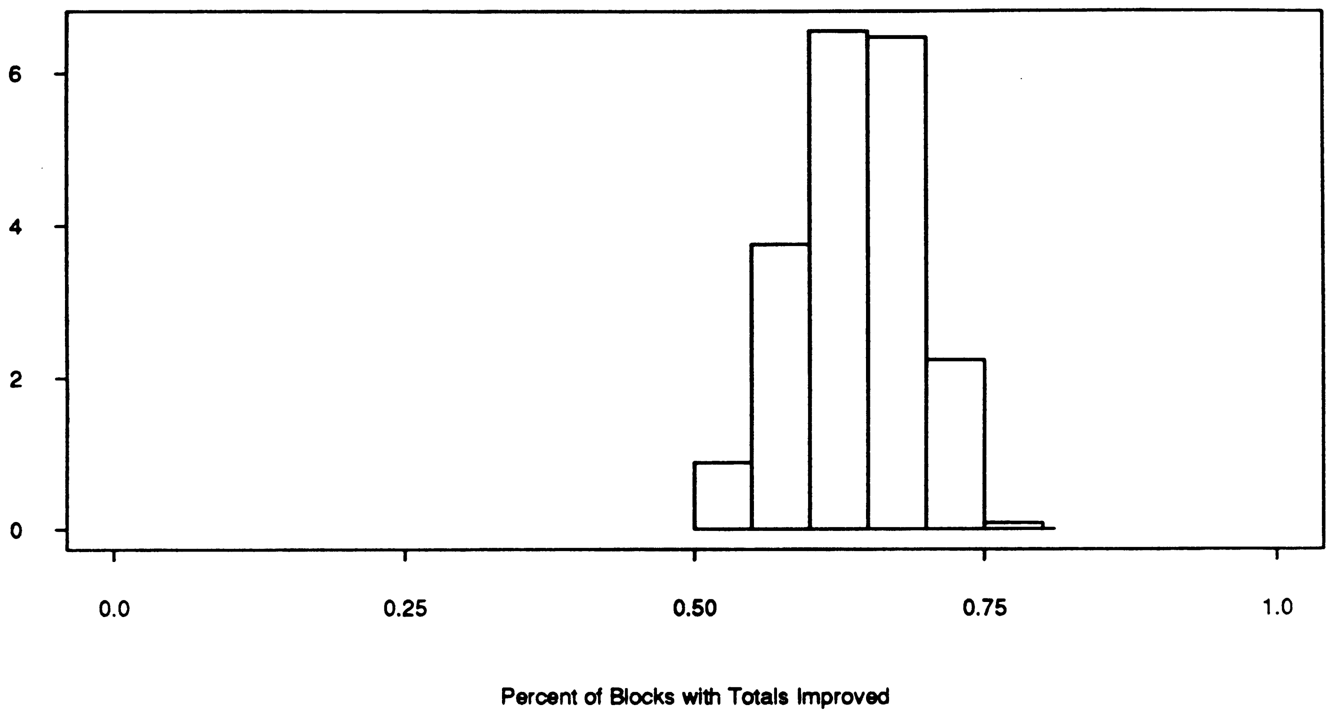


Figure A15: Percent of Blocks Improved, Arlington, Without Block Effects
(250 Simulations)



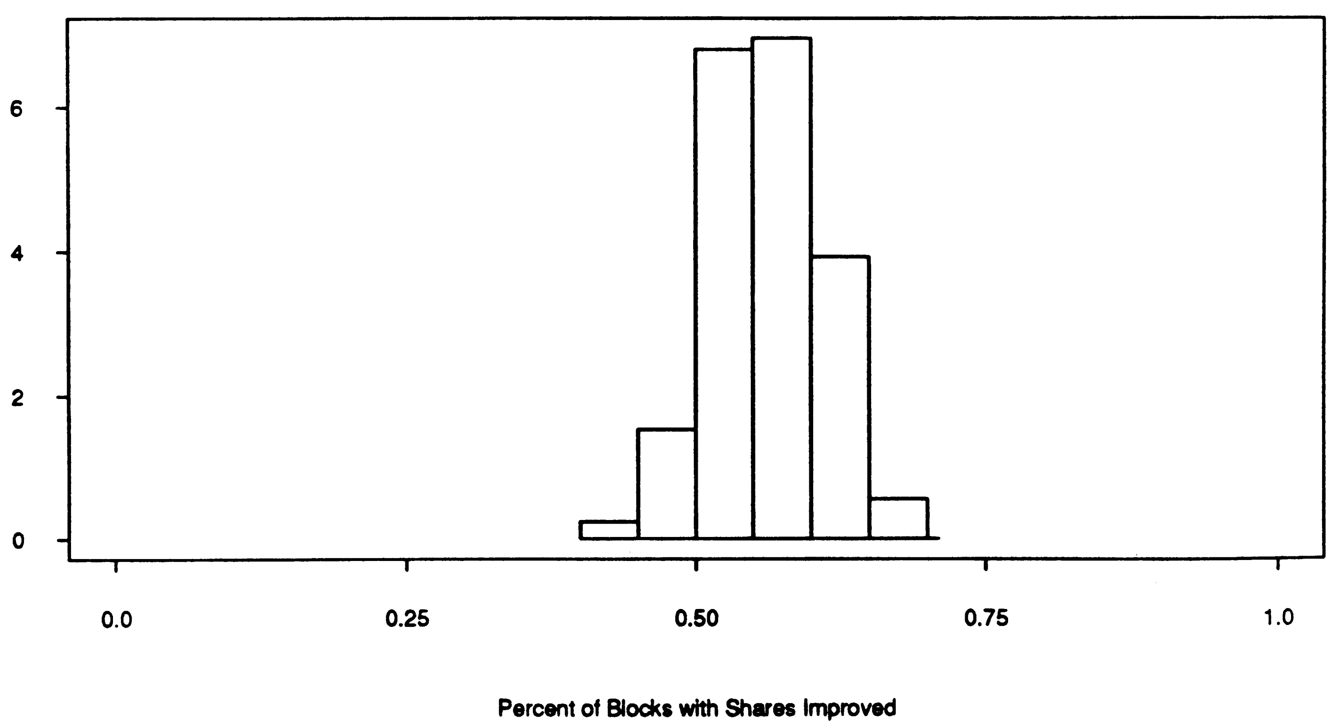
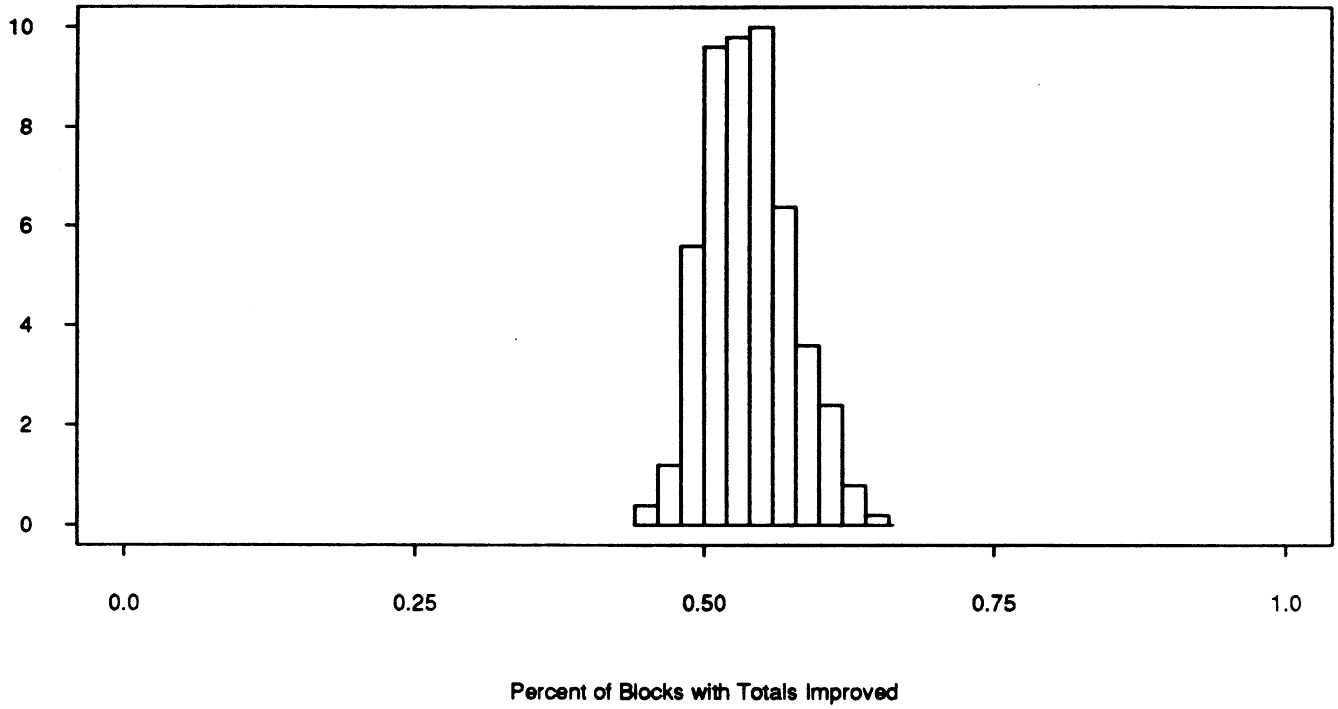


Figure A16: Percent of Blocks Improved, Arlington, With Effects + Int
(250 Simulations)



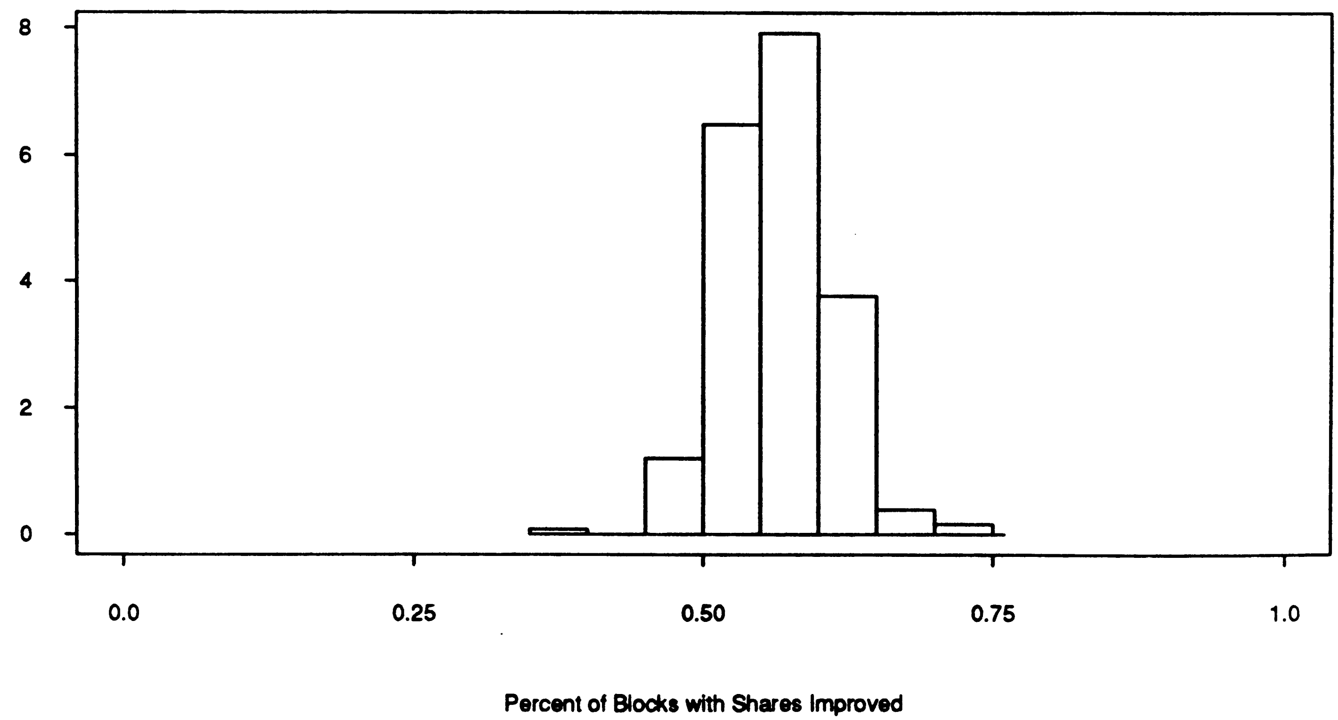
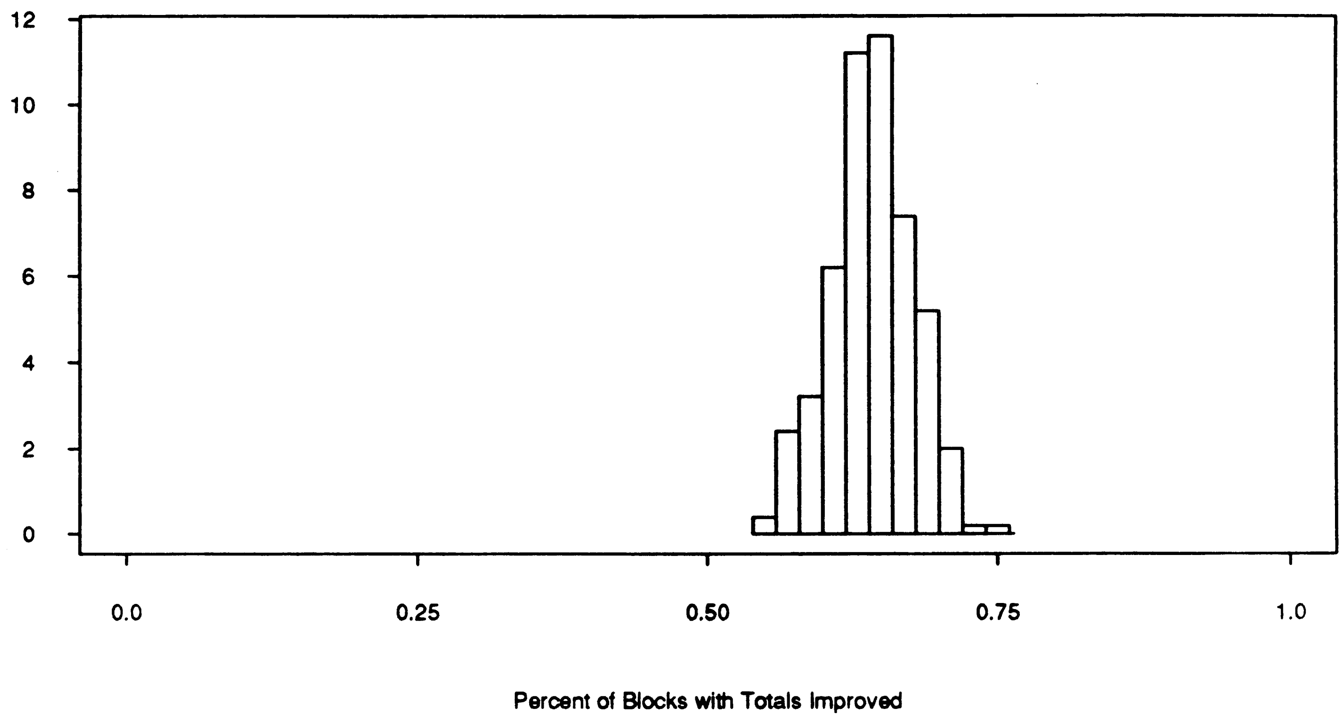


Figure A17: Percent of Blocks Improved, Detago, Without Block Effects
(250 Simulations)



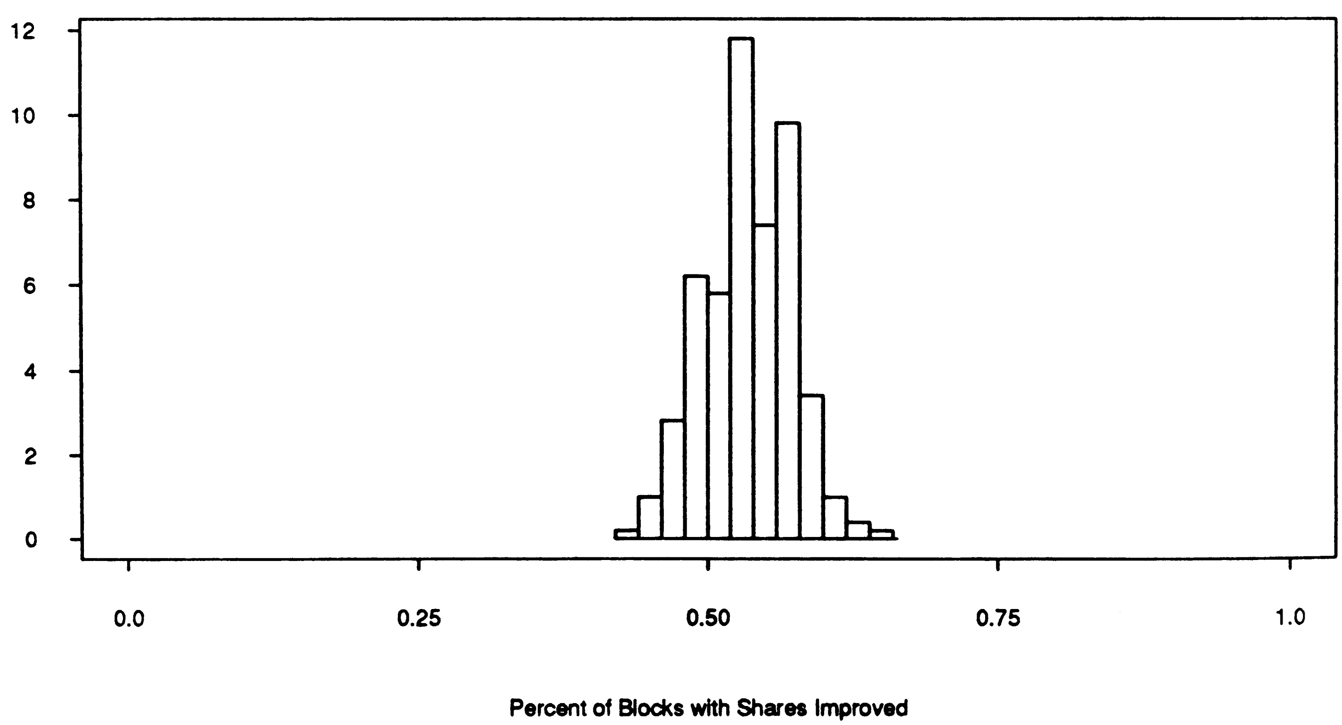
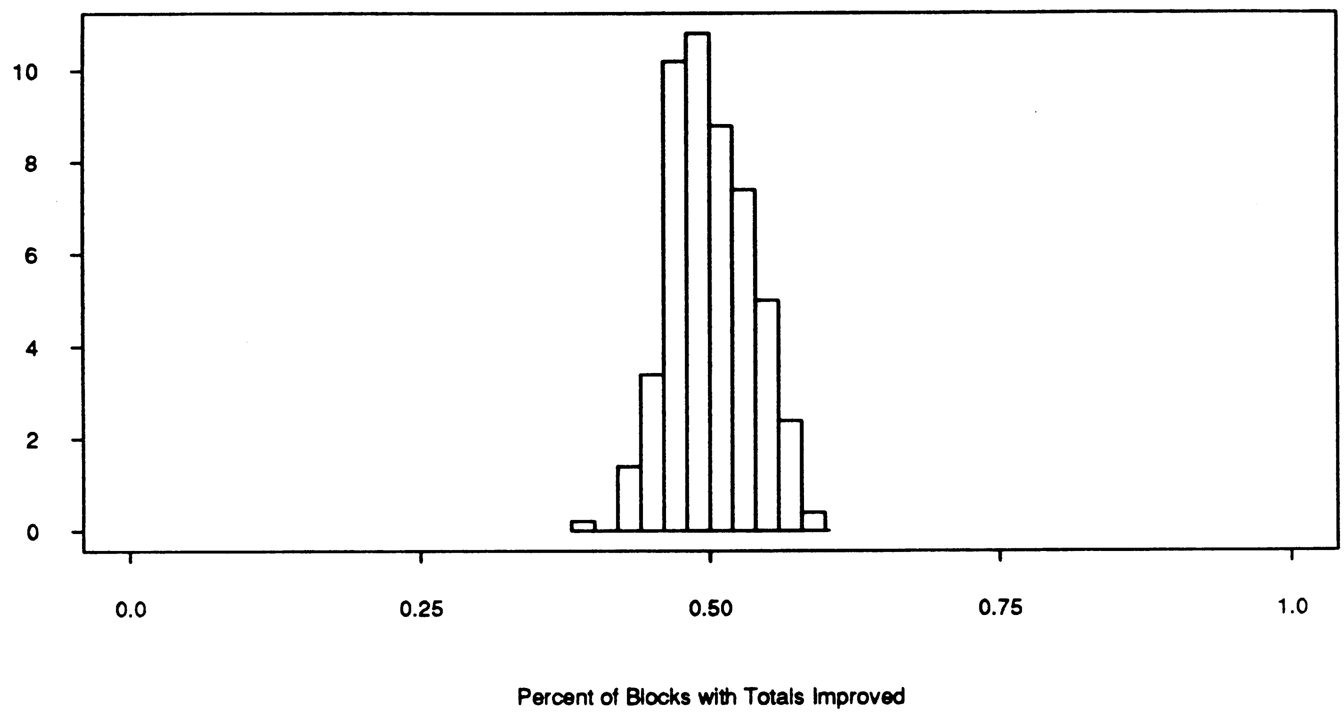


Figure A18: Percent of Blocks Improved, Detago, With Effects + Int
(250 Simulations)



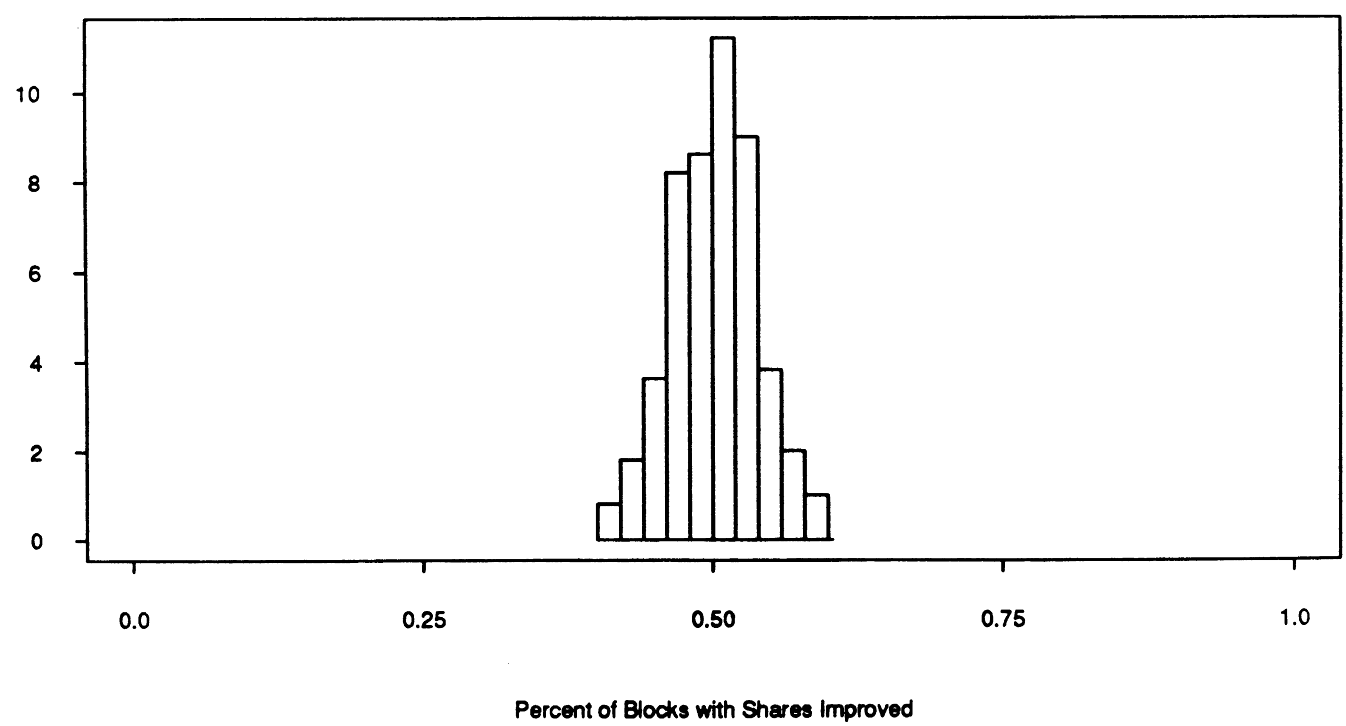
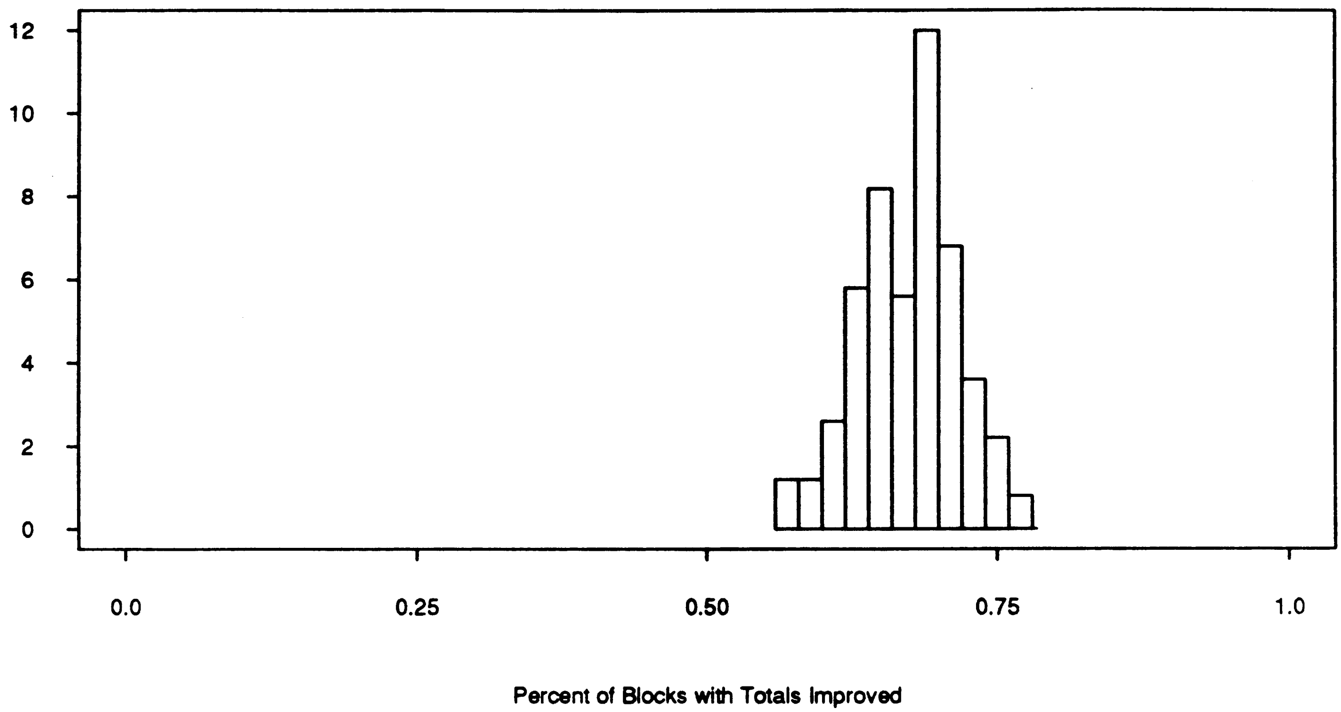


Figure A19: Percent of Blocks Improved, Berkeley, Without Block Effects
(250 Simulations)



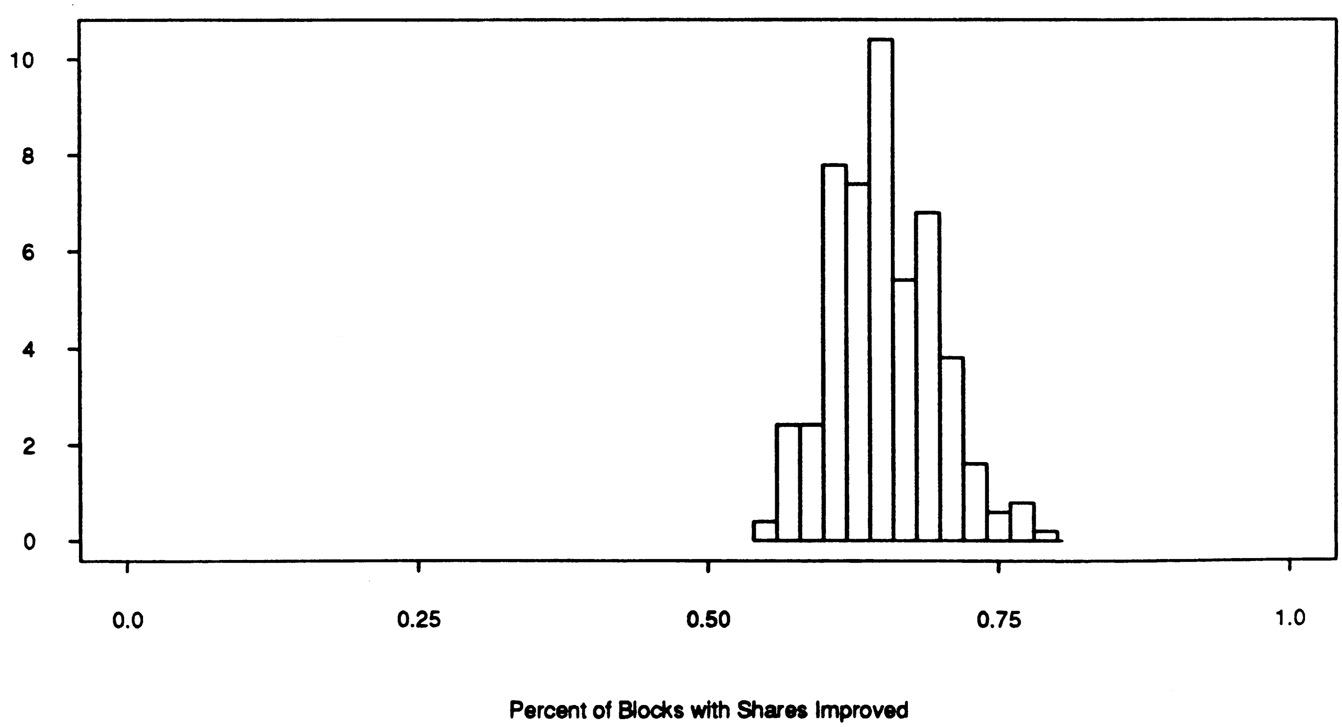


Figure A20: Percent of Blocks Improved, Berkeley, With Effects + Int
(250 Simulations)