Can Homelessness Be Prevented? Evidence from New York City's HomeBase Program

Peter Messeri, Mailman School of Public Health, Columbia University, New York

Brendan O'Flaherty, Department of Economics, Columbia University, New York

Sarena Goodman, Department of Economics, Columbia University, New York

December 16, 2011

Abstract: *In 2004, New York City established HomeBase in order to reduce the number of families entering its homeless shelters. Families who think they are in danger of becoming homeless can go to HomeBase offices to receive a wide variety of assistance, both financial and not, to keep them out of shelters. HomeBase started in different neighborhoods at different times. We use this variation in start-up to estimate the effect of HomeBase on shelter entries and exits. Our best estimates are that for every hundred families HomeBase enrolled, shelter entries fell by between 10 and 20. HomeBase had no discernible effect on the length of shelter spells. We believe that this is the first quasi-experimental evaluation of a homelessness prevention program.*

Acknowledgments: *We are grateful to the New York City Department of Homeless Services for data, financial assistance, and answers to our many questions; especially to Jay Bainbridge, Joanna Weissman, Eileen Lynch, Sara Zuiderveen, Jonathan Kwon, Veronica Neville, and Ellen Howard-Cooper. Scott Auwater and Bronxworks provided us with a detailed look at an operating program. Ingrid Ellen gave us lis pendens data. We are grateful for support from our institutional partners at CUNY, John Mollenkopf and Ellen Munley; and from the Columbia Center for Homelessness Prevention Studies, especially Carol Caton, Sue Marcus, Sue Barrow, Mireille Valbrun and Shoshana Vasheetz. We have received helpful advice from Serena Ng, Bernard Salanie, Beth Shinn and Till von Wachter. Maiko Yamagida, Abhishek Joshi provided excellent research assistance. The errors are own, as are the opinions expressed.*

Introduction

HomeBase (HB) is a community-based homelessness prevention program that the New York City Department of Homeless Services (DHS) has operated since 2004. The phased introduction of this program in different neighborhoods allows us to estimate the effect that this program has had on family shelter entries and shelter spells. HomeBase reduces shelter entries in a statistically significant fashion, and has practically no effect on shelter spells. There is great heterogeneity in the effects of HomeBase on shelter entries. One particularly strong and intuitive finding is that HomeBase is more effective at averting shelter entries in neighborhoods where there are a lot of entries to avert than in neighborhoods where there are few, if any, entries to avert.

To our knowledge, this is the first quasi-experimental evaluation of a homeless prevention program. See Apicello (2008) and Apicello et al. (2011) for reviews of the literature. Previous analyses of homelessness prevention programs in the literature have looked at whether participants have become homeless, but these analyses ignored several important counterfactual questions:

a. Would these participants have become homeless in the absence of prevention efforts?
b. When would these participants have become homeless in the absence of prevention efforts?
c. How many non-participants became homeless as a result of the prevention program—either because a fixed pool of housing, child care or employment training subsidies was diverted to participants, or because fewer unsubsidized vacancies arose, or because rents rose because of greater use, or because fewer child care or employment or treatment slots remained available?
d. When would non participants have become homeless?

e. How long would the participants who averted homelessness have stayed homeless in the absence of prevention efforts?

f. How long did the non-participants who became homeless because of the prevention program stay homeless?

g. How did the prevention program affect rates of shelter exit by people who were already homeless when the homelessness prevention program began?

While we cannot answer all these questions, and we cannot answer any of them in isolation, we can make good attempts at finding all the relevant sums, except those involving question g. By looking at the bottom lines of shelter entries and shelter exit hazards (rather than trying to follow individual families), we can come close to answering the question of how prevention programs affect homelessness.

Specifically, the change in entries that a prevention program causes is the number of participants who avoid homelessness (question a), minus the number of non-participants who become homeless as a result of the program (question c), minus the number of participants and non-participants whose entry was delayed (questions b and d). We estimate this sum, not the individual components.

Similarly, the change in exit hazard that a prevention program causes consists of three effects (questions e and f). We can estimate the sum of these effects.

We find that the net effect of HomeBase is to reduce shelter entries by between and 10 and 20 for every 100 HomeBase cases. HomeBase appears to have no effect on the hazard of shelter exits.

Because we have enough detailed data on the regular operations of the New York City family shelter system, we are able to construct plausible counterfactuals, and therefore we are able to come much closer to a complete analysis of the effects of a prevention program than previous work, which tried to answer one or at most two of these questions. However, because the counterfactual scenarios that are plausible span a considerable range, we cannot isolate a single unique number as "the" effect of HomeBase.

We estimated the effects of HomeBase operations by combining information from two administrative data sets. HomeBase client records allowed us to track the geographic distribution of HB services at the level of community districts and census tracts from its inception in November 2004 through the end of 2008. Family shelter records allowed us to track the number of families entering shelters and length of stay by community district and census tract between January 2003 and the end of 2008.

We estimated the effect of HB on shelter entries along two different dimensions. One dimension is geography: we use both community districts and census tracts as our geographic units of observation. Census tracts are more plentiful and exhibit greater variation in intensity of HB services, but community districts have cleaner definitions of when HB was operating and because they are larger than census tracts, allow us to net out more of any effects that HB might have in increasing entries by non-participants. Also, a larger percentage of census tract-months than community-district months have zero shelter entries, both before and after the introduction of HomeBase. The other dimension is how we quantify HB operations. In some equations, the key explanatory variable is a measure of HB **capacity**, and in other equations the key explanatory variable is the number of families that HB serves—HB **service**. The capacity equations are simpler and more direct but the great variation in the actual level of services that different offices provide means that they obscure a great deal of the heterogeneity in how many entries HB might divert in a month. The service equations resolve this heterogeneity issue but have to be estimated by instrumental variables techniques, and raise questions about the timing of HomeBase effects.

All of our estimates indicate that HB produces a statistically significant reduction in shelter entries in the neighborhoods in which it is operating. We also have considerable evidence that suggests that this reduction does not come at the expense of additional entries in nearby neighborhoods or in later months. HB appears to avert entries, at least on net, not divert or delay them.

Our estimates do not agree, however, on the size of the reduction in entries that HB causes. In terms of shelter entries averted per 100 HB cases, our capacity estimates range from about 12 to about 70. The service estimates range from 10 to 28 shelter entries averted per 100 HB cases. The differences in part arise from choice of unit of analysis. The CD level estimates cluster above 20 shelter entries averted per 100 HB cases. In contrast, the best CT level estimates cluster between 11 and 16. The differences also arise from the different counterfactual scenarios that these models implicitly work with. The estimates above the mid-20s generally rely on counterfactual scenarios that are implausible in subtle ways.

Reductions of even ten entries per hundred cases are large in the homelessness literature. Since homeless families on average stay in shelter for close to a year, these results imply a reduction in point-in-time (PIT) homelessness of at least almost ten (and maybe as much as twenty) per hundred HB cases. In contrast, subsidized housing probably reduces PIT homelessness by about 3-7 per hundred households served. (See Ellen and O'Flaherty 2010 for a review of the literature and calculation.) Of course, these results may not apply beyond New York, which has a constitutional right-to-shelter, and a large family shelter system, in which families stay for long periods.

We also look at the hazard of shelter exit for families who entered the system when HB during the study period. A priori, an effective prevention program might have conflicting effects on spell length. On one hand, it might disproportionately divert the families with the least serious problems and hence those who would have had the shortest spells; this would raise the average length of shelter spells. On the other hand, if HB delayed shelter entry for some families and these families ended their homelessness when their luck turned better, independent of what occurred in the shelter, then HB would shorten shelter spells. We find essentially no effect; HB appears to reduce the length of shelter spells but the effect is small and statistically insignificant.

The plan of the paper is the following. We begin with a history of HomeBase and a description of the data. The second section estimates the capacity effect, using both community-district and census-tract data, and the third section estimates the service effect. Section 4 addresses the questions of postponement and spillovers to non-participants. Section 5 analyzes exits and spell length.

1. Background

A. HomeBase history

For planning and land use purposes, New York City is divided into 59 community districts (CDs), each with about 135,000 people (the largest CD, Flushing, would have been the 70th largest city in the United States in 2000). The Department of Homeless Services (DHS) began HomeBase in November 2004 by selecting non-profit agencies in six CDs to operate the program. These CDs were not chosen randomly, although they were dispersed over four of New York City's five boroughs: DHS chose CDs that were heavily represented among the last addresses of shelter entrants. We will refer to these CDs as the "big six," because the majority of HB participants during the period we study came from these six CDs. These six included the CD with the most shelter entries during the 22 months before HB began, and even the big six CD with the smallest number of family shelter entries during the pre-HB period was ranked 14th. During the pre-program period, the average number of shelter entries from the big six CDs was over 2.3 times larger than the citywide average. Only residents of the big six CDs were eligible to receive HB services between the start of operations in November 2004 and June 2007.

DHS was pleased with the initial implementation, and expanded HB citywide in two phases. In July 2007, 31 more CDs were included in HB, and the remaining 22 CDs started in January 2008. The average CD that started in July 2007 had 21 percent fewer shelter entries in the pre-program period than the citywide average, and the average CD that started in January 2008 had 7 percent fewer. We will refer to these two sets of CDs as the "2007 cohort" and the "2008 cohort" respectively. After January 2008, HB operated everywhere in the city. In spring 2007, the initial contracts in the big six CDs were nearing completion, and because HB was going citywide, operators knew that the contracts for these particular CDs would be smaller. The intensity of service in the big six CDs declined in spring 2007, and remained low for the rest of the period. The addition of each successive cohort expanded the set of families eligible to receive HB services.

By "families" we mean any group of individuals who live together, and pregnant women. We include both families with children and families with no children or pregnant women ("adult families"). HB also served single adults unaccompanied by children, but we do not use data on this group.

HB was designed to help families overcome immediate problems and obstacles that could result in loss of housing. Families experiencing difficulties voluntarily apply at an HB office located in their neighborhood, or are referred from shelter intake centers (but only when an HB office is operating in the family's CD of origin). HB case managers have wide discretion in matching services to the specific problems of eligible families. Services include family and landlord mediation, legal assistance, short-term financial assistance, mental health and substance abuse services, child care, and job search assistance. Our data do not allow us to distinguish among services or service-delivery strategies. The neighborhood base of operations—the fact that each participating family has an address, if only tenuously—is the key to our analysis.

HB offices were instructed to provide services only to eligible families. Eligibility included both income and residence criteria. Although HB offices were supposed to serve only families living in their designated CDs, they did not always abide by this rule. About 5.8 percent of HB participants received services before HB was operating in the CDs in which they appeared to live. Presumably they travelled to an HB office outside the CD in which they lived. (It is also possible that they moved out of the CD while they were receiving HomeBase services. When a HomeBase participant moved, the address was written over.) We have rerun our results with the CD-months where these anomalies occurred excluded, and they do not change appreciably.

B. Data

We have data from January 2003 through December 2008.

Our primary dependent variables are family entries and exits in the New York City homeless shelter system, disaggregated by time and geography. We study only eligible families, who spent one or more days in the shelter system. Eligibility for the family shelter system essentially means a demonstration that a family has no reasonable alternative place to live. Determining eligibility often takes several weeks, and families who are found ineligible often re-apply. We do not have information on families who entered the shelter system and were subsequently found ineligible or who left before an eligibility determination was made. Homebase may have affected their number as well as the number of eligible entrants.

DHS supplied the team working on this project with a list of all eligible families who entered the shelter system between January 2003 and December 2008. Because information for December 2008 appeared to be incomplete, we excluded it from our analysis of entries and of HB operations. Information for each family included family composition, their entry date, their exit date (if it was before December 31, 2008), their type of exit, and their last address. Names were redacted. The Center for Urban Research at CUNY used its GIS capabilities to attach a community district and a census tract to each address, and to redact the address. We worked with a file that included community district and census tract[1], but no address.

The operation of HB provides us with several independent variables—some of them measures of capacity, some measures of service. The simplest measure of capacity comes from a listing of when HB began operation in each CD, and the exact location of each HB office. We also have a file of all HB participants from inception to December 2008. Like the shelter entrant file, this file has names and addresses redacted, but it includes the community district and census tract of the family's current address (not the office where they sought services). We also know eligibility status and date of enrollment. From this file we compute the number of families served by HB each month by the CDs and CTs they live in. (However, we are not able to tell whether the families that HB serves later enter shelters; Shinn and Greer (2011) examine this issue.)

Finally, the Furman Center at New York University provided us with a file of all lis pendens (LP) filings in New York City by date and address for the period 2001-2008. An LP filing is the first step in a foreclosure, although many LPs do not result in foreclosures[2]. After the LP is filed, the foreclosure and resale process often takes a long time—a year or more.

C. Summary statistics

---

[1] When a census tract spanned two CD's, for purposes of determining HB eligibility, we assigned it the CD with the larger number of cumulative shelter entries from that census tract.

[2] Of lis pendens filings in New York City in 2007, only 14 percent had ended with bank ownership or third party auction by 2009; 54 percent had had no subsequent legal transactions. See Furman Center 2010.

Table 1 presents summary statistics on the CD level. The average CD-month had about 10.8 shelter entries, and about 2.6 HB cases. But for many CD-months, HB was not operating and there were no HB cases.

Tables 2 and 3 present summary statistics on the CT level. The analysis covers 1,889 NYC census tracts that were either the last known residence of one or more families entering the City's shelter system between January 2003 and November 2008 or were the place of residence for one or more HB cases opened through November 2008. There are approximately 32 CTs for each CD and the much smaller size results in correspondingly smaller monthly shelter entries and HB cases opened. An average of .33 families entered the shelter system from a CT each month. During the period of HB operations, HB centers opened .11 cases during the average CT-Month.

A potential virtue of analyzing HB effects at the CT-level is the enhanced possibility of investigating temporal and spatial variation in the HB effects, because of the substantial increase in observations at much smaller geographic areas. We take advantage of the CT data to investigate two groupings of census tracts. First, we investigate whether the effect of HB depends on a CT's use of shelter space. We would like to know if HB works differently in neighborhoods with many potential shelter entrants than it does in neighborhoods with few.

To measure use of shelter space, we assume the census tract characteristics that determine risk of homelessness, the supply of affordable rental apartments, socioeconomic status of residents, family and employment stability are relatively constant during the six-year period of this study. Therefore, we measured a census tract's stable use for family shelter space during the operation of the HB program, as the count of families entering the shelter system with the census tract as last residence for the 22 months preceding the start of HB services: January 2003 through October 2004. Consistent with this assumption, the correlation between the count of families entering the shelter system by census tract for the 22 months just prior to the start of HB services and first four years of HB operations is .93. Even eliminating the many "low use census tracts", those with only 0, 1, or 2 families entering shelters during the pre program period, the association dropped imperceptibly to .90 for 894 census tracts with between 3 and 67 family entries during the pre-program period. For purposes of subgroup analysis, census tracts are grouped into low use CT's (0, 1, or 2 cases opened during the pre program period), moderate use CT's (3 to 18 families) and high use CT's (19 to 67 families).

Table 2 summarizes the numbers of census tracts and monthly observations available for these strata. The large majority of study census tracts, 995, fell into the lowest shelter use group. There are 702 moderate use and 192 high use census tracts. The first 32 months of HB operations (through June 2007) were officially restricted to residents of the 233 census tracts located in the six original CD's. During the 22-month, pre-HB period, 2,748 families or 23% of all family shelter entries came from these census tracts. Starting in July 2007, an additional 914 CT's became eligible for HB services, and the final 742 CT's became eligible when the program went citywide in January 2008.

The intended restriction of HB services in its early years to neighborhoods that were the source of large numbers of shelter entries is clearly evident in Table 2. Whereas high use CT's represented only 10% of all study CT's, they constituted 26% of the original 233 census tracts in the Big Six. By contrast, low use CT's represented 53% of all study CT's, but made up only 15% of the initial group of HB eligible CT's.

Table 3 presents information on the distribution of HB cases and shelter entries. The values in each cell are for the specified CT grouping and time period (1) the number of shelter entries, (2) HB cases opened, and (3) the average number of cases opened per CT-month--a measure of service intensity. For the number of CT's and CT-months refer to the corresponding cells in Table 2. During the time period under study, 10,314 HB cases were opened during the time CT's were officially eligible for services. During this time period 16,800 families entered the shelter system from these CT's. These cases were opened during 35,117 CT-months (see Table 2) for an average of 0.29 cases opened monthly in a service eligible CT. A small number of cases, 631, were opened in CT's prior to the start of their official eligibility. The contamination of HB exposure in non-eligible CT's was negligible, .011 cases per CT-Month or 1/30 the average service intensity for eligible CT's.

Over the course of the study period and for various CT subgroups, there was substantial variation in both the intensity of HB services as measured by the average number of cases opened per month in a CT and the ratio of HB cases opened to shelter family entries. During the restricted phase of the HB program operations (November 2004 through June 2007), the average number of HB cases opened each month in a service eligible CT was .85. During this early period, the total number of HB cases opened in eligible CT's, 6,370, exceeded the number of families entering shelters from these CT's, 4,345. Following the program's expansion in July 2007, a substantial dilution in the intensity of HB services is clearly evident. During the last 17 months of the study period, the total number of HB cases opened, 3,944, was substantially smaller than the number of family shelter entries, 12,455. There was a corresponding dramatic drop from .85 to .14 in the mean number of HB cases opened each month in service eligible CT's.

As expected, HB services are targeted to CT's with highest use of shelter space. Although a greater number of HB cases were opened in the more numerous moderate use service-eligible CTs than high use CT's (5,070 vs. 4,628), the mean number of HB cases opened each month in moderate use CT's, .34, was about a third of the average cases per CT-month in high use CT's, 1.96. Turning to the much more numerous low use CT's we find that the absolute number of HB cases (616) and intensity of services (.040 cases per CT-month) is considerably smaller than that for either the moderate or high use CT's.

D. Trends in shelter entries

HomeBase began operations during an extraordinarily dynamic period in the use of the City's family shelter system. During the study period, 2003-2008, an average of 635 families entered the shelter system each month, but the mean obscures pronounced seasonal and year-to-year fluctuations (see Figure 1). In any calendar year, entries tend to peak in the late summer months and ebb during the winter and early spring. The amplitude of seasonal variation is large, on the order of a 200 to 300 change in family monthly entries. The start of HomeBase in November 2004 occurred midway through a period of relatively low shelter use that extended through 2005: entries fluctuated between 400 and 600 each month. After 2005 family entries began to climb. The numbers of monthly family entries increased to between 600 and 800 during 2006, 2007 and the first half 2008. By the second half of 2008, monthly family entries jumped to over 1,000 as the burst of the housing bubble and the full force of the Great Recession began to take hold.

Of particular note for evaluating HomeBase is the sharp downward disjuncture in family entries that followed the start of the program in November 2004. Family entries attained a study period nadir for the seven months following the start of HomeBase operations in the big six CDs. The sharp fall-off in entries is in part due to seasonal declines, but a regression model that adjusts for both seasonal variation and annual change in family entries estimated this disjuncture to be a highly significant -75 (95% C.I.= -122,- 27) drop in monthly shelter entries averaged over 50 months of partial and later full coverage HomeBase operations . It is not possible to draw strong causal inferences about the effects of HomeBase from a citywide aggregated time series, because of the possibility that changes in economic and housing market factors that coincided with the start of HomeBase operations were also influencing the observed secular changes in shelter entries. For a more robust estimate of treatment effects, we turn to more detailed analyses at the CD and CT levels. Figure 1 also illustrates the very "noisy" background in shelter entries that presents significant challenges in modeling the counterfactual condition that would have obtained in the absence of the HB program.

2. Shelter entries: Capacity analysis

A. Methods

<u>CD level</u>

In this section we treat HB operation as a series of binary variables, and see how changes in shelter entries in CDs subject to the "treatment" of HB operation differ from changes in untreated CDs. Specifically, the simplest equation is:

$$S_{ct} = \alpha + \beta H_{ct} + \gamma_c + \delta_t + \varepsilon_{ct}. \qquad (1)$$

Here c indexes CDs and t indexes months, S_{ct} denotes the number of shelter entrants from CD c in month t, γ_c is a CD fixed effect, and δ_t is a month fixed effect. The key independent variable is H_{ct}, a dummy variable equal to one if and only if HomeBase is officially operating in CD c in month t. A negative β indicates that the treatment works: HomeBase reduces shelter entries. The coefficient β is an estimate of the average number of shelter entries averted in a CD-month of HB operation.

Simple equation (1) can be improved in several ways.

First, foreclosures are month- and CD-specific events that may affect shelter entries, since most households affected by foreclosure in New York City were probably renters (Furman Center 2010) or may reflect CD-specific housing market trends. Let F_{ct} denote the number of LP filings in CD c in month t. We add several lags of this variable to equation (1): contemporaneous, 3-month lag, 9-month lag, 12-

month lag, 15-month lag, and 18-month lag. We use these lags because the time between the filing of an LP and the resolution of the foreclosure, including the displacement of residents, is often long.

Second, the official definition of when HB was operating in a CD does not properly account for treatment, since some families received HB services before HB was operating in their CD, as we noted in the discussion of the small number of apparently ineligible families receiving services. To address this problem, we add a second dummy variable P_{ct}, equal to one if and only if some resident of CD c has received HB services during or before month t. During normal operations of HB, both H_{ct} and P_{ct} equal one, and so we will be interested in the sum of their coefficients.

Third, treatment may change as time goes on. One possibility is that HB offices become more proficient as they acquire more experience. In that case, later months of experience would be associated with greater reductions in shelter entries. On the other hand, HB may delay shelter entries rather than avert them entirely. Then the first months of operation would show the greatest reduction in total entries; after that, delayed entries would offset new reductions. Participants who come to HomeBase in later months when the program is better known may also differ systematically from participants who made their way to HomeBase when it was little known. Another possibility is that as time passes, enthusiasm wanes and treatment intensity declines. To explore these possibilities, we include a dummy variable R_{ct}, equal to one if and only if in month t HB has been officially operating in CD c for more than two months.

To summarize, our fullest model of entries as a function of HB capacity is:

$$S_{ct} = \alpha + \beta_1 H_{ct} + \beta_2 P_{ct} + \beta_3 R_{ct} + \gamma_c + \delta_t + \sum_s \varphi_s F_{c(t-s)} + \varepsilon_{ct}. \qquad (2)$$

Of course, we also estimate less complete equations.

Equation (2) is linear. While the functional form of the estimating equation does not matter for the argument that the coefficients on HB operation give average reductions in shelter entries associated with HB, it does matter for the interpretation of the fixed effects and hence the construction of the counterfactual (what would have happened if HB were not operating in the CD-months in which it was operating?) . In particular, the form of equations (1) and (2) forces the month fixed effect to be the same in absolute magnitude in every CD: a city-wide shock like a change in rules or a recession increases or decreases shelter entries in every CD by the same amount.

We address this problem in two different ways, neither of which is totally satisfactory. One approach is to estimate equation (2) in logarithmic form:

$$\ln(S_{ct} + 1) = \alpha + \beta_1 H_{ct} + \beta_2 P_{ct} + \beta_3 R_{ct} + \gamma_c + \delta_t + \sum_s \varphi_s F_{c(t-s)} + \varepsilon_{ct}. \qquad (2')$$

The implicit counterfactual in equation (2') is that city-wide shocks operate in multiplicative fashion: they cause identical percentage changes in shelter entries in every CD. The possible drawback of equation (2') is that it forces entries averted to be a constant fraction of entries not averted. By contrast, linear equation (2) is less restrictive about the relationship between entries averted and entries not averted. Equation (2) provides a better model of HB operations if those operations are small relative to flows into the shelter system: if HB offices run out of resources to treat cases, for instance, not cases to treat, or good cases to treat. If, for instance, an HB office is incapable of averting more than a small fixed number of cases in a month, no matter how many families are entering the shelter system, then the coefficient on HB operations in (2') will be biased down.

An alternative way to construct the counterfactual is to stratify the sample and estimate equation (2) separately for CDs that usually have many entries and CDs that usually have few. This allows us to estimate different sets of month effects for the CDs that are usually large and for the CDs that are usually small. Operationally, we divided the sample into two strata: the 21 CDs that had more than the mean number of shelter entries in the pre-program period covered by our sample, and the 38 CDs that had fewer than the mean number of entries. Stratifying the sample and estimating two equations has an additional advantage in interpretation. Although we do not attempt to perform a cost-benefit analysis, anyone who wanted to use our results for this purpose could be guided by the stratified results in deciding whether to expand or contract the program in high use or low use CDs.

Stratifying the sample and thereby multiplying the number of month fixed effects, however, has a downside. The month fixed effects are supposed to reflect city-wide shocks, but if they are not constrained to be equal or proportional across the two strata, they may not reflect the same shocks at all. Adding sets of month fixed effects almost automatically reduces the size and significance of the CD-month-specific coefficients like those on HB operations and foreclosures. In the limit, if there were 59 strata so that each CD had its own set of month fixed effects the effects of HB and foreclosures would be unidentified. Thus when we stratify, we will have to examine the resulting month fixed effects and see how well the month effects from one equation are correlated with those from the other—essentially, whether they are picking up the same city-wide shocks.

Including 18-month lags of foreclosures forces us to ignore 18 months of data, starting in 2003. We have estimated these equations over the full period without foreclosure lags (and for the short period without foreclosure lags). Results are not materially different. We have also estimated these equations dropping the CD-months in which unofficial entries appear. Again, the results are not materially different.

<u>CT level</u>

Except for adjustments related to the change in unit analysis, the CT level equations that estimate the average treatment effect of the HB program are identical to CD level capacity equations (1), (2) and (2'). Corresponding to the change in unit of analysis, the dependent variable is now monthly counts of family shelter entries for each census tract and CT fixed effects substitute for CD fixed effects. All models also include monthly fixed effects. The HB capacity variables are identical to those applied to the CD level analysis. That is to say, all CT's are assigned their CD's values for official and unofficial start of HB services.

The effect of foreclosures on shelter entries over an 18 month period is also estimated at the CT level. In a minor departure from the CD level analysis, all monthly lags from 1 to 18 months are estimated in the CT equations. Because of their much smaller size, monthly LP housing unit filings are relatively rare occurrences at the CT level. There were no LP filings in three-quarters of monthly CT observations and 1 to 5 LP filings accounted for another 23 percent of CT-months. At the other extreme, 52 monthly observations with the largest numbers of LP filings ranged between 59 and 690 units. Despite their very small number (.04 percent of all observations), preliminary analysis indicated that these extremely high LP counts had a substantial depressing effect on estimates of family entries. To remove the distorting effect of very high LP counts, the LP count was truncated at 50 and an indicator variable for these outlying observations was added for each lagged variable.

The large number of CTs allows additional examination of heterogeneity. The basic capacity model in equation (2) is augmented with interaction terms between each of the three capacity measures and dummy variables for moderate (DMOD) and high (DHI) use census tracts.

$$S_{ct} = \alpha + \beta_1 H_{ct} + \beta_2 P_{ct} + \beta_3 R_{ct} + \beta_4 H_{ct} * DMOD_{ct} + \beta_5 P_{ct} * DMOD_{ct} + \beta_6 R_{ct} * DMOD_{ct} +$$

$$\beta_7 H_{ct} * DHI_{ct} + \beta_8 P_{ct} * DHI_{ct} + \beta_9 R_{ct} * DHI_{ct} + \gamma_c + \delta_t + \varepsilon_{ct} \qquad (3)$$

A second grouping of census tracts is based on the time when each CT became officially eligible for HB services, in one of November 2004, July 2007 or January 2008. We speculate that HomeBase program effects may vary by cohort because expansions of HB services bring in new subset s of census tracts that differ in homelessness risk characteristics. The Big Six CD's were intentionally selected because of their unusually high rates of shelter entries. The eligibility cohorts also differ in the level of HomeBase services availability and possibly the skill of HomeBase centers to resolve housing problems. The geographic expansion of HomeBase services also corresponded to a diminution of resources available to each center after June 2007, as funding was now distributed over larger number of offices serving a greatly expanded geographic area. Finally the operation of HB in each successive cohort of census tracts spanned different secular trends in family shelter entries. Citywide family shelter entries remained at relatively low levels during the period when the HomeBase program was restricted to the Big Six CD's. Coincident with the expansion of the HomeBase program after July 2007, there was an acceleration in the numbers of families entering the shelter system that continued through the end of the study period in November 2008. The capacity equations to estimate the separate effect for each wave of census tracts are similar to (2) and (3)except that the subgroup dummy variables now indicate census tracts that became formally eligible for services in July 2007 and January 2008.

As in the CD-level analysis, the CT-level capacity equations are estimated in both linear additive and log linear forms. We fit the log linear form to a Poisson regression model that is most appropriate when the outcome takes the form of count data. While there were good reasons to prefer a linear additive model when modeling HB operations at the CD-level, there are good theoretical reasons for preferring the log-linear or Poisson model when fitting CT data. In contrast to the linear additive model, the Poisson regression coefficients are ratios that measure proportionate change in the outcome variable associated with a change in the predictor variable. A Poisson coefficient of .9 for HB the presence of HB operations can be interpreted as a 10% reduction in shelter entries. Although HB capacity is essentially fixed at the CD level, the rationale for the linear additive model, the Poisson model allows for the possibility that HB resources and hence the number of HB cases are allocated proportional to the need of CT's for homeless prevention services, as indexed by the number of families entering the shelter system. Our theoretical reasoning is consistent with the empirical distribution of HomeBase cases shown in Table 3. Low use census tracts averaged .040 new HB cases opened each month during period of official program eligibility. The mean number of new HB cases per months jumps to 0.34 and .96 for moderate and high use CT's, respectively. DHS reinforced the natural tendency of HB to target high-risk families from areas where shelter use is greatest by contracting services to nonprofit organizations located in or close to neighborhoods with the highest volume of family shelter entries. Of course even if HB offices allocate resources roughly proportionate to CT-level need for services, this doesn't necessarily translate into equal effectiveness at the individual family level. We test this assumption when HB effects are estimated separately by CT's level of use for shelter space.

If there is concern that the proportionality assumption may not hold--e.g. if the absolute size of the true HB effect is similar in high and low use CT's similar, then a linear model may be preferred. Since the linear model makes less stringent uses on the data, it may result in a more precise estimate of the average treatment effect across all CT's, even if the program effect is proportionate to shelter entries. However, when the effects are proportional, the linear model may not be as accurate as the Poisson model in predicting shelter diversions at the CT level or for groupings of CT's that differ by frequency of shelter entries. Because the linear model predicts the same amount of change for all CT's, the model parameters will predict too many family diversions in CT's with low shelter use for shelter and too few diversions for the high use CT's.[3] A conclusive head-to-head empirical test of the linear versus the log-linear specifications is not possible with the data at hand, but separate estimates of the HB for each of the three use strata may generate some insights into the plausibility of the alternative

[3] In formal terms, when treatment effects are proportional to the size of the outcome variable, the linear additive estimate may produce an unbiased average treatment effect, but there will be substantial effect heterogeneity. The Poisson model fit to the same data will result in a ratio estimate of the average treatment effect that is much more homogenous.

assumptions of the two models.

Effects of foreclosure are estimated for the linear model, but not the log-linear model, since each additional foreclosure is assumed to have a fixed additive effect on shelter entries regardless of spatial aggregation or the volume of families entering the shelter system.

B. Results

HB appears to decrease entries to the shelter system. This holds at both the CD and CT levels.

<u>CD level</u>

Table 4 for the CD level shows the results from OLS regressions like (2) that use HB capacity measures as the independent variables.

HB operation appears to reduce shelter entries by about four per CD-month. Out of 1,063 CD-months of official operation, 118 assumed a value of zero for the experienced variable. Hence the average effect of a CD-month of official operation was a reduction of 4.57 shelter entries. Since the average CD-month during our study period had about 10.8 shelter entries, this reduction is economically as well as statistically significant. The sign of the coefficient on experience indicates that HB offices become less effective (although not significantly so) after they have been open for two months; this suggests that delay is probably playing a role, though a small one (i.e. some potential entries are occurring later rather than not at all).

Since the average operating HB office served 10.49 families (some of whom were from outside the CD) in an average month and averted 4.57 entries, one way of interpreting this result is to say that on average 43.6 shelter entries were averted for every 100 families that HB served. This does not imply that the number of shelter entries averted would be proportional to the number of HB families served if more or fewer families had been served.

While foreclosures are not directly our concern, we note that they increase homelessness. These are the first results we are aware of that attempt to link foreclosures to homelessness. The coefficients indicate that for every 100 LP filings, about five additional families enter shelters. The effect is strongest at 12 months after the initial filing. Since most LP filings do not result in foreclosures, the results are economically as well as statistically significant. However, we have not attempted to test for causality.

Results for the logarithmic equation (2') are not so strong. They are presented in table 5. Coefficients have the same signs as they do in table 4, but are less often significant (except for experience, which is not significant in either equation). The estimated magnitude of the effect of HB is also somewhat smaller, but is still significantly different from zero when all the coefficients are added. Since the average CD-month has about 10.8 entries, HB is producing a reduction of around 1.58 entries in the average fully operating CD-month, rather than over four with the linear specification.

The estimated effect of foreclosures is also smaller and less significant. This suggests that the effect we are seeing is from displaced tenants directly, not from a general indication of housing market difficulties.

The linear model appears to fit the data much better than the logarithmic model, but the fixed effects of the linear model run into the problems we discussed in the methods section. Consider CD 503, the south shore of Staten Island, a relatively affluent CD. During the pre-program period, less than half a family entered the shelter system from this CD in an average month. Yet the counterfactual with the linear model is that absent HB, around 9 families would have entered in August 2008).

Because of this problem, we present the results of the stratified equations in table 6. To save space, we present the results only with all three HB variables. HB effects are quite small and insignificant for the stratum with small CDs, and while they are significant for the stratum with large CDs, the effect is smaller than the unstratified effect. Specifically, on average the stratified coefficients imply that a fully operating CD averts 1.72 shelter entries a month, while the unstratified coefficients imply 4.59. (The 95 percent confidence intervals do not overlap.) The effect of foreclosures in the stratified equations is also smaller: 100 LP filings lead to around three shelter entries, as compared to five with the unstratified equation. This effect is highly significant, however.

The correlation in month fixed effects between the two stratified equations is 0.65, considerably less than one, and so part of the reason for the attenuation of the HB and foreclosure effects is the mechanical addition of new fixed effects, as we argued in the methodology subsection. But the correlation is positive and of considerable size, and so stratification is providing some useful information.

Figure 2 shows smoothed monthly fixed effects for the unstratified equation, and for the two strata. The unstratified monthly fixed effects are flat for most of the period, and rise strongly at the end, probably largely because of the beginning of the Great Recession (Cragg and O'Flaherty 1999 and O'Flaherty and Wu 2006 show that New York City family shelter entries rose in the previous two recessions). The story that the unstratified equation tells is that the fall in large CD entries at the start of the period was due largely to HB (which was then operating only in six of the 21 large CDs), and that shelter entries would have risen even more sharply in 2008 without HB. The counterfactual story that the stratified equations imply is different. At the beginning of the period, entries would have fallen sharply in the large CDs without HB—essentially a fall in entries from the 15 large CDs without HB in the beginning plays a prominent role in constructing the counterfactual here, while it is swamped by the 38 small CDs in the unstratified model. Fixed effects for the large CDs rise as sharply as they fell after the end of 2005, but they barely return to the initial level by the end of 2008: for the large CDs, the recession is seen as offsetting the earlier fall, not raising entry rates above 2004 levels. Relative to 2004, the recession effect is mainly felt in the small CDs, not in the large.

Allowing the data to find two different patterns of month fixed effects thus tells a fundamentally different counterfactual story from the story that results from forcing month effects to reflect a single pattern of city-wide shocks. In the unified story, conditions were even until the recession hit, and when it hit, deteriorated severely. In the other counterfactual, the unified story explains what happened in the city's more prosperous neighborhoods, but not in the poorer neighborhoods. Those neighborhoods saw conditions improve considerably until the end of 2005, and then slide back to 2004 levels. The "tale of two cities" produces smaller estimates of the effects of HB and of foreclosures than the tale of a unified city. The improvement in the poor neighborhoods but not the more prosperous is not implausible—many neighborhoods may produce practically no shelter entries in good times or bad, and so improving conditions at the beginning of the period may have had almost no impact on them.

(We also experimented with interacting month fixed effects with preprogram shelter entries in the linear model. This produced small and insignificant estimates of CD effects, especially of unofficial operation. Because DHS selected the Big Six for their large number of shelter entries in the pre-program period, and because CDs with large numbers of shelter entries are more likely to generate early starts of unofficial operations, these equations have collinearity problems. CT level results were similar.)

<u>CT level</u>

Table 7 presents estimates of the effect of HB capacity on shelter family entries for the Poisson version, and Table 8 presents capacity model estimates using an additive linear form.

Following the start of formal operations of HB services, the Poisson model yields a marginally significant decline in family entries (IRR=.967, 95% CI= [.931,1.005]). The effect of HB capacity across all CT's and eligible time periods was equivalent to averting 3.3 family shelter entries for each hundred families entering the shelter system (not for each hundred families that HB saw). When extrapolated to the entire period of HB operations, this model estimates that 573 shelter entries were averted during the 49 months of official HB operations. When the model parameters are expanded to include variables for unofficial start of HB services and more than months of official operations, the estimated average HB program effect increases to .947 (95% C.I.=.881,1.018) or a diversion of 5.5 families for each 100 shelter entries. For the entire study period, this model estimates that HB services averted 1,244 entries.

The linear formulation of the capacity model estimates stronger HB program effects. For this model, the average treatment effect of HB during periods, when CTs were officially eligible for HB services, was -.071 entries per CT-month. Extrapolating across all CT-months of HB operations results in an estimate of 2,493 families averted from shelters. When additional variables are added for unofficial operations and experience, an average treatment effect of -.13 is obtained for official HB operations after two months of operations. When extrapolated for the entire period of operations, the total number of shelter entries averted because of HB services is estimated to be 7,184.

Estimates of HB effect for CT's that differ by historical use of shelter space and the timing of the official start of HB operations suggest that HB effects are heterogeneous, but the estimates across models do not necessarily produce a consistent pattern. The Poisson and linear capacity models produced qualitatively similar patterns when HB effects are estimated by level of shelter use. Focusing on the fully specified capacity model, both the Poisson and linear model indicate that the HB effect strengthens with increasing use for shelter services. The Poisson model suggests that for low use census tract HB services may actually induce shelter entries, whereas the linear model suggests the HB has no effect when serving low use census tracts. In contrast to the overall average effects, the separate estimates by level of use are strengthened for the Poisson model and weakened in the linear model. The Poisson model contain only one set of city-wide monthly fixed effects, but the strata specific linear models estimate monthly effects for each stratum.

When we examine census tracts based on timing of official eligibility for services, a consistent pattern occurs across all three models. HB services have their strongest effects in census tracts that first became eligible for services starting in July 2007. The estimated HB effects are much weaker or absent for the Big Six and the 2008 cohort. The models are not consistent with regard to the size of the effect for these latter two groups of CT's. The Poisson model suggests a modest effect for the Big Six and no effect for the 2008 cohort, but the linear model suggests the reverse pattern. Again, these models have only one set of monthly fixed effects.

With stratified linear equations—that is, with a different set of monthly fixed effects for each stratum—the results on average are closer to the Poisson model than to the linear model.

Figure 3 shows the monthly fixed effects from the unstratified and stratified linear models. The correlations between the monthly fixed effects of the different CT strata are considerably higher than the correlation between different CD strata with CTs: between .80 and .94, as opposed to .66 for the CD strata. Unlike the stratified CD month effects, the stratified CT month effects tell a consistent story: conditions are stable at the beginning of the period and get worse at the end. The month effects at the end of the period are also larger for the highest use CTs than the other CTs; this is also a difference from the CD results.

C. Reconciliation of capacity results

Table 9 consolidates all of the capacity results, and presents them on the consistent basis of entries averted per 100 HB cases. We use historical accounting in this table: the number of months in each condition (unofficial operation, official operation, and experienced operation) is that which occurred in the study period, not that which DHS might anticipate going forward. The results consistently show that HB reduces shelter entries, but are not consistent on the size of the reduction.

Generally, CD models imply larger effects than CT models. For models with only one set of monthly fixed effects, however, the CD and CT results have the same pattern. For an unstratified linear model both data sets imply a very large reduction in entries , about 68 entries per hundred HB cases for the CD model and about 65 for the CT model. For the unstratified logarithmic or Poisson model, both data sets imply smaller reductions: about 29 entries per hundred cases for CDs and 12 per hundred for CTs. The difference between the CD and CT logarithmic estimates, however, is not statistically significant. For both data sets, the stratified linear model produces very similar results that are much closer to the unstratified logarithmic or Poisson models: 21 entries per hundred cases for both the CDs & CTs.

Since we have reason to believe that the unstratified linear results are too high, we are left with a range of 10 to 20 shelter entries averted per hundred HB cases. Note that this is an historical average. If HB had concentrated only on neighborhoods with heavy shelter use, it would have done considerably better. If it had concentrated only on neighborhoods with light shelter use, it would have done considerably worse.

3. Shelter entries: Service analysis

A. Methods

CD level

An alternative approach is to assume that what affects shelter entries is not the simple presence of an HB office, but the number of families that the HB office serves. Both direct and indirect effects of

HB depend on how many families are served. This relationship argues for making the independent variable the level of HB service in a month, not a description of the capacity for service.

Let HB_{ct} denote the number of families living in CD *c* who were served by HB in month *t.* Thus we would like to estimate a linear equation like:

$$S_{ct} = \alpha + \beta HB_{ct} + \gamma_c + \delta_t + \sum_s \varphi_s F_{c(t-s)} + \varepsilon_{ct}. \qquad (4a)$$

or a quadratic equation like

$$S_{ct} = \alpha + \beta HB_{ct} + \check{\beta} HB_{ct}^2 + \gamma_c + \delta_t + \sum_s \varphi_s F_{c(t-s)} + \varepsilon_{ct}. \qquad (4b)$$

However, HB service in a particular CD-month is likely to be endogenous—for instance, a CD-specific event that causes many families to visit HB is likely to cause many other families to enter shelters at the same time. So estimating equation (4) by OLS will not produce unbiased estimates.

Fortunately, administrative decisions provide us with several instruments for HB service, and so we estimate equation (4) by instrumental variables. The first instrument is obvious: H_{ct}, the variable indicating whether HomeBase was officially operating. CD-months when HomeBase was officially operating should see more HB cases than CD-months when it was not. The next instrument is the distance from the CD in question to the HB office that was nearest that month. This is zero for CD-months when HomeBase is officially operating.

We also use a series of instrumental variables to reflect administrative differences between groups of CDs and differences over time. For each of the three cohorts defined by starting date (the big six, the 2007cohort, the 2008 cohort) we have a dummy variable that turns on only for the CDs in this cohort, and only after the start of official HomeBase operations for that cohort. In addition, for each cohort and each fiscal year after HomeBase operation begins for that cohort, we have a dummy variable that turns on for CDs in that cohort in that fiscal year. We use fiscal year dummies because most DHS contracts are on a fiscal year basis and many important policy changes coincide with new contracts, including changes in funding levels that influence the overall number of HB cases the HB centers are able to open each year.[4]

Thus, for instance, for a CD-month when HomeBase is not operating, all instruments will be zero, except the distance to the nearest operating HomeBase office.[5] For a CD-month in fiscal year 2009 for a CD in the 2007 cohort, the positive dummy variables will be H_{ct}, the 2007 cohort dummy, and the FY-2009-interacted-with-2007-cohort dummy.

These instruments are highly correlated with HB service; the R^2 on the first-stage CD-level regression for HB_{ct} is .797. The distance variable and H_{ct} are significant and their estimated effects have the right sign.

The other question about the instruments is whether they satisfy the necessary exclusion restriction. To be valid instruments, they must affect shelter entries only through the services that HB provided. Unconditionally the location of HB offices and the period of formal HB operations in a CD are correlated with number of shelter entries. Thus the big six CDs that were first to receive HB services were chosen because of their high rate of shelter entry. The expansion of HB services and the opening of new HB centers to serve the remaining CD occurred during the final 18 months of the study period when the rate of shelter entry was higher. However the inclusion of the month- and CD- fixed effects should remove the correlation with shelter entries mediated by the level of HB services. There remains the possibility that DHS may have responded to transitory events in a particular month and CD, but we

[4] Of course, we omit dummies as appropriate.

[5] In contrast to the capacity model in which observations prior to the start of HB operations in November 2004 provide useful information, observations for the service model are restricted to the period of HB operations.

regard such a fine-tuned response as highly unlikely. DHS created HB capacity through contracts with nonprofit organizations, and nonprofits had to gear up and acquire staff and space to implement the contracts. These are complex and time-consuming processes. HB was implemented in only three waves, moreover, so there was little scope for adjusting start-up dates. Hence it is likely that the effect of CD-and-month-specific events on HB capacity was at most small. To the extent it was present, however, it biases down our estimates of the effect of HB service on shelter entries.

Because of the size of our sample and because all of the big six CDs are in the large stratum, we were unable to produce meaningful stratified results.

CT level

The CT service model is estimated using the linear form and instrumental variables applied to the CD-level data. However, we do not estimate a quadratic form of the service model, as the observed number of HB cases opened in any month at the CT level is seldom more than 3 cases. The declining effect of HB with increased volume is better understood as CD-level mechanism that operates as a contextual factor in a CT-level analysis. In this study, we did not attempt to investigate the potential conditioning effect of this and other CD level attributes on CT level effects.

For the CT service equation, we also obtained 2SLS estimates separately for each use stratum.

B. Results

CD level

Table 10 shows the results of the IV regressions, both the linear and quadratic versions. Both equations imply that HB averts shelter entries, but the size of the effect is considerably larger, on average in the quadratic equation than in the linear.

In both regressions the sum of foreclosure coefficients is positive and significant—slightly larger in fact, than it was in the linear OLS regressions. In the first-stage regression, moreover, foreclosures also increase HB cases. This suggests that our decision to use IV was not misguided: if foreclosures raise both shelter entries and HB cases, other CD-and-month-specific shocks probably do the same. (In fact, if equation (4a) is fit by OLS, the coefficient on HB cases is a positive and significant: HB appears to increase shelter entries.)

In the linear specification, it appears that shelter entries fall by about 10.3 for every 100 families that HB serves. This value is consistent with the CT capacity results with the logarithmic specification.

The quadratic specification nests the linear, and the significant coefficient on the quadratic term indicates that the linear equation is mis-specified. The quadratic results imply that the marginal effectiveness of an HB office declines considerably as the number of families it serves increases. Taken literally, the quadratic results imply that after about 30 cases, more cases are counter-productive (but very few CD-months have more than 30 cases). Based on the quadratic specification, an HB office of average size, seeing about 10.33 cases a month, would reduce entries by about 4.9 a month, which is in line with the linear capacity estimate.

The more relevant comparison, however, is the average reduction in entries which, by Jensen's inequality, is less than the reduction in the average CD-month. Several possible counterfactuals can be explored, but they all give roughly the same answer. For comparison with CT results, the relevant comparison is all CD-months. The average CD-month had 2.62 HB families served, and the average squared number of HB cases was 76.71. This implies an average entry reduction of 0.75, or 28.6 shelter entries averted per 100 HB cases. (Alternatively, it implies that HB averted a total of 3137 shelter entries during our study period.) Going forward, the question is how many shelter entries can be averted in months with HB officially operating. Among CD-months with HB officially operating, the average number of cases from within the CD was 9.703 and the average number of squared cases was 296.986. This implies an average reduction in shelter entries of 2.645—or 27.26 entries averted per 100 HB cases. (Restricting either set of CDs to those in the IV regression sample does not materially change these results.)

The service results also imply a larger impact of foreclosures on shelter entries than the capacity results imply. In the long run, a hundred LP filings result in 8.2 more shelter entries. This estimate is higher than the estimate from the capacity equation, and most other estimates of foreclosure effects in this paper do not support such a large impact.

Why is HB more effective with small caseloads than with large? We do not know. Answering this question is clearly important for future decisions about the size of HB.

One possibility is simple congestion. With staff fixed, participants in CD-months with higher caseloads will receive less attention and outcomes may deteriorate as a result.

Selection is another possible reason for the differential effectiveness. The size of the coefficients on P_{ct} suggests that selection is an important part of the story. These coefficients indicate that unofficial operation of HB in a CD reduces shelter entries by about two a month. Unofficial operation without official operation ("pure unofficial operation") may mean that residents of that CD have travelled outside of it to receive HB services, and have been served—probably contrary to regulations. During an average CD-month of pure unofficial operation, 0.58 families were served (many CD-months of pure unofficial operation had zero families because unofficial operation starts when the first family in the CD receives services). The astounding effectiveness of unofficial operation is probably an important part of the diminishing marginal effectiveness result.

Why is pure unofficial operation so effective? Families, who travel outside their neighborhoods, may be special, especially those who can convince HB workers to bend the rules. Either the families themselves or the HB workers may recognize that they are in imminent danger of homelessness and that specific assistance can resolve that danger. The same may be true for HB offices that are small and not well-known. Convenient, well-publicized HB offices may draw many families that HB cannot help avoid homelessness.

<u>CT level</u>

Table 11 presents 2SLS estimates for the CT services equation. The service model indicates that each 100 HB cases translates into a reduction of 12 in shelter entries or approximately one averted entry for every 8 HB cases opened. This extrapolates to 1,313 entries averted that results from all HB cases opened during the study period. This is approximately the value we extrapolated for the CT capacity model, and about half the number of diversions estimated from the CD equations.

Heterogeneity also appears to be important with service equations, but the pattern is not the same as the pattern with capacity equations. In contrast to the capacity models, a very strong effect is associated with each case opened for families living in low use census tracts and there are at best modest effects for cases opened in moderate and high use CTs. This seems consistent with CD quadratic service results.

Taken together the results for the stratified CT equations capacity and service models suggest that low use census tracts have few families seeking HB services, but a higher probability that a family seeking services will be diverted from shelter entry.

For census tracts sorted by eligibility cohort, the largest effects are for the cohort that became eligible in July 2007. The 2004 cohort has a smaller but still statistically significant negative effect.

Unexpectedly the HB effect associated with the January 2008 cohort is positive, indicating that with an increasing number of cases opened the number of entries also increases. We have no explanation for this counterintuitive finding. This result might be related to a subtle specification error for the instrumental variables, but we can't determine its precise nature. Reconciliation of service results

Table 9 also includes a summary of the service results.

With the service models, the CD and the CT linear models produce about the same estimate: 10 to 12 averted shelter entries per hundred HB cases. The CD quadratic estimated over twice this level, or 26 cases averted.

We should not be surprised that the CT linear and CD linear service estimates are about the same. Our instruments are at the CD level, not the CT level; they use no CT-specific information. Thus we are looking at the variation in the number of CT HomeBase cases that is driven by CD-level phenomena. By construction, month-to-month variation in HB cases is constrained to be the same in each CT of a CD (indeed, the same in each CT of a cohort of CDs).

Month-to-month variation is shelter entries in a CD is the sum of month-to-month variation in shelter entries in the component CTs, and so if month fixed effects are working about the same way on the CD and CT levels, the estimated effect of HB cases on shelter entries should be the same in the CT linear and CD linear equations. The CT linear and CD linear estimates will also differ because the number of contained CTs differs between CDs. But differences between CT linear and CD linear estimates require more explanation than similarities.

Thus the approximate equivalence of the CT linear and CD linear estimates gives us confidence in the service equations. The CD quadratic equation nests the CD linear equation, and lets us reject the hypothesis of no non-nonlinear effects. The nonlinear effects we find on the CD level have no obvious counterpart on the CT level. (Congestion, for instance, arises from an imbalance between the resources of a CD level office and the total number of families seeking service, not matter how they are divided among CTs.) Thus the CD quadratic is our preferred service estimate.

4. Issues and possible overstatements

The effect of HomeBase on shelter entries might be overstated for two reasons: because we may not have properly accounted for the effect of HomeBase on non-participants, and because HomeBase may merely delay shelter entry rather than precluding it. In this section we discuss these two issues. Our estimates in the previous two sections, especially the capacity equations, appear to be net of both effects, but a great deal still needs to be learned.

A. "Musical chairs" and other effects on non-participants

Strictly speaking, we do not know whether the shelter entries that were averted would have been HomeBase participants or non-participants. On a CT level, our estimate is *net* shelter entries averted: the gross number of entries averted minus the number of entries caused by HomeBase. It is plausible that the former category consists mainly of participants and the latter of non-participants, but we do not know. Nor do we know the size of the gross effect—only the net effect. But the net effect is the relevant number for cost-benefit analysis.

If HomeBase service to residents of one census tract made non-participants in nearby census tracts more likely to enter shelters, then our CD estimates of the net effect of HomeBase would be lower than our CT estimates. They are not. Thus the net effect of HomeBase service in a census tract on non-participants in nearby census tracts appears to be zero.

Another possibility is that HomeBase service in a CD could affect non-participants in nearby CDs adversely. To check for these geographic spillovers, we formed pairs of adjacent CDs, and looked at entries from each pair for each month. (Since there are an odd number of CDs and Staten Island with three CDs is relatively isolated, we consolidated all of Staten Island as "CD-pair.") [6] If HomeBase service in a CD affects non-participants in adjacent CDs adversely, then the estimated HomeBase effect estimated from CD-pairs will be smaller than the effect estimated from single CDs. We estimated only the capacity equation because the preferred service equation is nonlinear. Table 12 shows the result. The HB capacity effect is larger when estimated on double-CDs, not smaller. Thus, we find no evidence of spillovers between adjacent CDs.

There still could be larger, more diffuse spillovers—HomeBase activity in Queens may affect non-participants in Staten Island. We have no way of checking for such spillovers. Since we have found no evidence of inter-CT intra-CD spillovers or of spillovers between adjacent CDs, we do not think spillovers between non-adjacent CDs are likely to be large.

---

[6] Specifically, outside of Staten Island, we first constructed an adjacency table describing which pairs of CDs were adjacent. We consider the East River (but not the Harlem River or Newtown Creek) to be impenetrable (and so CDs on opposite sides of the East River are not adjacent, but CDs on opposite sides of the Harlem River are). (The Harlem River is treated differently from the East River because it is bridged more often.) Similarly we considered Central, Van Cortland, and Flushing Meadow Parks to be impenetrable, but not the Bronx Zoo or Forest Park.

We then arranged the boroughs in the order Manhattan, Bronx, Brooklyn, Queens; within boroughs CDs are numbered. We then did iterations of the following process:

a. Check to see whether any unassigned CD has only one potential unassigned partner; if so, assign the CD and that partner.
b. When no unassigned CD has only one potential partner, assign the unassigned CD that comes first in order to the unassigned adjacent CD that comes first in order.
c. Start over at a.

Clearly this algorithm is not unique.

Thus the estimates in the previous sections appear to approximate the net HomeBase effect. The gross HomeBase effect on participants may be larger, and could be found with a randomized controlled experiment. Such an experiment, however, could not find the net effect. By combining the results from a randomized experiment with the results from this paper, one might be able to learn whether HomeBase affected non-participants within a census tract.

B. Postponement

HomeBase has effects beyond the month in which service starts, and these effects may alter the interpretation of our results, particularly the service equations. It is helpful to begin with a precise understanding of intertemporal linkages.

Let V(t) denote the probability that a family receiving HomeBase services remains out of the shelters for t months after these services begin, t = 1, 2,…. Define V(0) = 1. Let v(t)= V(t-1))-V(t) denote the probability that a family enters shelter in month t, conditional on receiving HomeBase services. Let U(t) be the probability that this same family would remain out of shelter for t months if it did not receive HomeBase services—the counterfactual. Set U(0)=1 and let u(t) = U(t-1)-U(t) denote the corresponding probability of entering shelter for the first time in month t. Let q(t) = u(t)-v(t), t =1, 2, ….

In the long run, the expected number of shelter entries that one HomeBase case averts is the limit of U(T)-V(T) as T goes to infinity. Define

$$E(\infty) = \lim_{T\to\infty}[U(T) - V(T)] = \lim_{T\to\infty}\sum_{t=1}^{T} q(t)$$

This is what we would like to know. The naïve estimate of HomeBase effects is q(1), shelter entries averted in the month service begins.

A priori, there is no reason to think that E(∞) is either bigger or smaller than q(1). Some families whom HomeBase serves may enter shelters in later months; they might have entered in later months without service. Some HomeBase families without HomeBase might have entered shelters in later months; because of HomeBase they never enter. The expression

$$E(\infty) - q(1) = \lim_{T \to \infty} \sum_{t=2}^{T} q(t)$$

can be either positive or negative.

Let HB(t) denote the number of HomeBase cases begun in month t of calendar time. If HomeBase starts operating in month 0, then the net reduction in entries through month T due to HomeBase is

$$E^{*}(T) = \sum_{t=0}^{T-1} HB(t)q(T-t)$$

The most direct approach is to try to estimate the sequence (q(t)). Since the CT service equation is plausibly linear, we estimate shelter entries from a CT as a function of current HomeBase cases and lags of HomeBase cases, with lags extending up to six months. Table 11 shows the result. The sum of the contemporaneous effect and six lags results is substantially larger: -.33 compared with the contemporaneous estimate of -.11. We are not sure what to make of the large increase in the lag effects for the instrumented number of cases. However it surely argues against a postponement hypothesis.

We undertook a second analysis that is parallel to the one applied at the CD level to the geographical question. We re-estimated both the capacity and services model by now grouping observations into longer time periods, 2, 3 and 6 months. The estimates in table 13 show only modest declines in the HB effect when the duration of the observation period is lengthened up to six months,

Thus we cannot reject the hypothesis that the sum of q(t) for t running from 2 to 6 is zero. This suggests that ignoring lags does not lead us seriously astray in the service equations. If the sum of u(t)-v(t) for t running from 7 to infinity is zero, too, then q(1) =E(∞). This result is not "no postponement"; it is "no net postponement", or "equal postponement regardless of HomeBase services."

With a quadratic term and less data, lags are not practical with CDs.

The interpretation of the capacity equations is simpler. The capacity effects without the "experienced" variable are unbiased estimates of the average value of E*(T) during the study period. With the experienced variable included, the estimates with R=0 give the average value of E*(1) and E*(2), and those with R=1 give the average value of E*(T), $T \geq 2$. Thus the capacity coefficients are estimates of the HomeBase effects that are uncontaminated by postponement problems; but they cannot be confidently extrapolated to the future without a more serious structural model like those in the service equations. However, the approximate equivalence of the estimated HomeBase effect in the capacity and services equations is also weak evidence for "no net postponement."

A final weak piece of evidence for "small net postponement" comes from the study of residuals from the capacity equation for CDs. If short-term net postponement occurs, then months with unusually large numbers of shelter entries averted should be followed by months with unusually low numbers of entries averted, as some of the families served in the first month enter in the second. So the serial correlation between residuals should be more negative with HomeBase is operating than when it is not operating.

To test for such an effect, we use our fullest model of entries as a function of HB capacity, equation (2). We divide the original sample into two subsamples. In one subsample of CD-months, HB services are “on”, and in the other they are “off.” We compute three sets of residuals from equation (2) estimated over our samples (the original, the “on” subsample, and the “off” subsample), and then estimate a linear regression of each residual set on its one-period lag. Note that a t-test using the estimated coefficient on the lagged residual in any of these three equations is a test for residual autocorrelation at lag one (assuming all regressors are strictly exogenous). However, in our case, we are primarily interested in whether the estimable serial correlation is statistically significantly different between our two subsamples.

To answer this, we conduct a Chow Test. The Chow statistic is computed using the sum of squared residuals from each lagged equation, the number of observations in each subsample (N1 and N2), and the total number of parameters (k); in our case, the test statistic equals 5.66. The test statistic follows the F distribution with k and N1 + N2 − 2k degrees of freedom, in general, and 3 and3121 in our case. The null hypothesis for this test is that the regression fit is equal across our three proposed samples, or in other words that the serial correlation is the same for the full sample and both subsamples.

We find that the serial correlation is more positive when HB is operating than when it is not. When HB is officially operating, serial correlation between residuals is 0.894 (s.e.=0.015), and when HB is not officially operating, the serial correlation is 0.825 (s.e.= 0.013). The study of residuals provides no evidence to support the existence of net postponement.

The critical value is 1.091 at the 5% significance level (with a p-value of 0.0007), so we reject the null hypothesis, concluding that there is strong evidence that the expected values in the three groups differ. The residuals are more-positively correlated when HB operates, which is not consistent with the postponement story. But the effect, even if significant, is small.

Thus our capacity results are unbiased estimates of the average effect of HomeBase operation, and several pieces of evidence suggest that net postponement is either small or nonexistent. An experiment would be able to shed light on this issue because it will be able to estimate the gross effects of HomeBase, U(t) and V(t), not just the net effects, and will be able to account for longer lags and more complex patterns. The experiment could show either more or less postponement.

5. Exits and spell length

   A. Theory

To find the full effect of HB on New York City's family shelter population, we must look at exits as well as entries. HB could affect exits in three different ways. Two of these ways would make spells longer, and one would make spells shorter.

Selection might make spells longer. If HB were more successful in averting homelessness for families with less serious problems than for families with more serious problems, and if homeless spells are longer for families with more serious problems, then HB will be more successful in averting spells that would have been short than in averting spells that would have been long. The average spell that starts when HB is operating would therefore be longer than the average spell that starts when HB is not operating.

Spillovers might also make spells longer. For instance, if HB participants stay in apartments they would otherwise vacate, fewer apartments will be available for shelter residents to move into. When HB is operating, then, spells may end less often.

On the other hand, postponement could make spells shorter. Suppose some families leave shelters when an exogenous favorable event occurs—winning the lottery, finding a good job, getting married—and the hazard of the good event does not depend on whether the family is in a shelter or not, and rises over time. Then if HB delays shelter entry, it reduces the expected interval between shelter entry and the favorable event. The average shelter spell for families who entered after HB started would be shorter.

Notice that each of these three effects tells us to look at a somewhat different set of exit hazards. Postponement tells us to look at all days in spells that began after HB had been operating a few months. Selection tells us to look at those days, as well as all days in all the other spells that began when HB was operating, too. The spillover story, by contrast, tells us to look at days when HB was operating, not the spells in which they are embedded.

To keep the analysis simple, we will concentrate on selection and postponement, and not test directly for spillovers. To the extent that spillovers are geographically diffuse, moreover, our dataset may not let us test for them at all (we have no way of knowing whether a family that originally came from Brooklyn would not have left the shelter system and moved to the Bronx if HB had not been operating).

The majority (about two-thirds) of eligible families leave the shelter system by receiving a subsidized apartment through DHS. This is called "placement." Families' circumstances and decisions have some bearing on when placement occurs, but DHS resources, rules, and queues are very important. We would not expect selection to have a large effect on the timing of placement, and postponement should have none, since time-in-shelter is what matters for DHS queues.

B. Methods

We use Cox proportional hazard methods. Because we want to consider placement and non-placement exits separately, we use a competing risk specification. The goal is to find whether families who entered the system when HB was operating had a smaller or larger hazard of non-placement exit than other families.

Specifically, let $\lambda_j(d, f, m, c)$ denote the hazard that in calendar month *m*, family *f*, which came from community district *c*, will leave the shelter system after *d* days, and the exit will be type *j* (placement or non-placement). Our basic equation is

$$\lambda_j(d, f, m, c) = \lambda_0^j(d) \exp\{[\beta X_{fc} + \gamma_c + \delta_m] + L[\tilde{\beta} X_{fc} + \tilde{\gamma_c} + \widetilde{\delta_m}]\} + \varepsilon_{jfcdm}$$

(5)

Here $\lambda_0^j(d)$ is the baseline hazard for type j exits; the Cox method does not estimate this directly. The vector X_{fc} is a vector of characteristics of family *f* and community district *c*. In particular

$$\beta X_{fc} = \beta_1 A_f + \beta_2 K_f + \beta_3 \overline{H_{fc}} + \beta_4 \overline{P_{fc}} + \beta_5 \overline{R_{fc}}.$$

Here A_f is the (demeaned) number of adults in family *f*, K_f is the (demeaned) number of children in family *f*, $\overline{H_{fc}}$ is a dummy equal to one if and only if H_{ct}=1 for the month in which family *f* entered the shelter system, $\overline{P_{fc}}$ is a dummy equal to one if and only if P_{ct}=1 for the month in which family *f* entered the shelter system, and $\overline{R_{fc}}$ is a dummy equal to one if and only if R_{ct}=1 for the month in which family *f* entered the shelter system.

Continuing with (5), γ_c and δ_m are dummies for the CD and the current month respectively, and *L* is a dummy variable equal to one if and only if the exit is a placement. Thus the independent variables are fully interacted with placement type.

The coefficients we are most interested in are $\beta_3, \beta_4, \beta_5$. These coefficients indicate how the non-placement hazard changes for families who entered the shelter system when and where HB was operating. We are also interested in $(\beta_3 + \widetilde{\beta_3}), (\beta_4 + \widetilde{\beta_4})$, and $(\beta_5 + \widetilde{\beta_5})$. These sums indicate how the placement hazard changes for families who entered the shelter system when and where HB was operating.

C. Results

Figure 4 shows the baseline survival probabilities for placement and non-placement exits for families who entered when HB was officially operating and those who entered when it was not. (This is for a family with the mean number of adults and children.) HB appears to make little difference, especially for placements.

Table 14 provides the results from estimating equation (4). It confirms the general picture that HB makes no significant difference to exits. HB families seem to leave slightly faster, but the difference is tiny and not statistically significant. Family size has large and significant effects: larger families stay longer, though the effect is stronger for non-placements than for placements. In fact. families with more children are placed sooner than families with fewer children; the effect is small but significant.

Selection and delay postponement may both be operating, but if they are, they are cancelling each other out. There is reason to suspect, however, that neither is operating. If operating in the hypothesized direction, postponement would increase the non-placement hazard rate for CDs that have been operating several months, relative to CDs that have been operating less time. So postponement implies that the effect of experience on non-placement exits should be positive and large. Selection has no implication for a change in effect after several months. So if both postponement and selection are strong and offsetting, the coefficient on experience for non-placements should be positive and significant. But it is not: the point estimate is .004 (z= .09).

Notice that neither selection nor postponement would affect the length of shelter spells in a Markovian world where current condition was a sufficient statistic for all predictions. This result is weak support for a Markovian modeling of shelter transitions, but it may also indicate that net postponement is not large.

6. Conclusion

Our best estimates are that HomeBase reduced shelter entries by between one and two for every ten families it served. It did not change homeless spell lengths. Our evidence suggests that, on net, non-participants did not enter shelters more often as a result of HomeBase, and that either "small net postponement" or "no net postponement" occurred. We cannot perform a benefit-cost analysis because we do not know how much HomeBase cost, or what the value of averting a shelter entry is.

While we have been able to estimate net effects, especially historical net effects, we have not been able to estimate gross effects—particularly gross postponement. Experiments would be very helpful in estimating many of these gross effects.

HomeBase effects are heterogeneous, and we have not been able to figure out whether that heterogeneity is due to different activities by service providers, different characteristics and circumstances that HomeBase participants bring to HomeBase offices, or different environments. We are pretty sure that HomeBase worked better in neighborhoods that historically generated many shelter entries than in neighborhoods that historically generated few. This is intuitive—you can't avert shelter entries that would not have occurred anyway. Even in neighborhoods where potential shelter entries were plentiful, however, HomeBase had decreasing returns to scale: it was less effective when it had to contend with many cases. Whether this deterioration is due to congestion or to selection is an open and important question. Small operations in areas of great need seem to have worked best.

References

Apicello, Jocelyn, 2008, “Applying the population and high-risk framework to preventing homelessness: A review of the literature and practice,” working paper, Columbia Center on Homelessness Prevention Studies.

---, William McAllister, and Brendan O’Flaherty, 2011, “Homelessness: Prevention” forthcoming, article 00352 Encyclopedia of Housing and Home.

Cragg, Michael, and Brendan O’Flaherty, 1999, “Do homeless shelter conditions determine shelter population? The case of the Dinkins deluge,” Journal of Urban Economics 46: 377-415.

Ellen, Ingrid Gould and Brendan O’Flaherty, 2010, Introduction, in Ellen and O’Flaherty, eds., How to House the Homeless. New York: Russell Sage.

Furman Center for Real Estate and Urban Policy, 2010, Foreclosed Properties in New York City: A Look at the Last 15 Years. Accessed on January 21. 2011 at

http://furmancenter.org/files/publications/Furman_Center_Fact_Sheet_on_REO_Properties.pdf.

O’Flaherty, Brendan 2009, “Homelessness in the United States,” Oxford University Press Handbook of Planning and Urban Economics.

--- and Ting Wu, 2006, “Fewer subsidized exits and a recession: How New York City’s family homeless shelter population became immense,” Journal of Housing Economics 15 (2): 99-125.

Shinn, Mary Beth, and Andrew L. Greer, 2011, “Report to Department of Homeless Services and Abt Associates on Using HomeBase Data to Predict Shelter Entry and Length of Stay Among Applicants for HomeBase Services,” mimeo.

Table 1
Summary Statistics: CD Level

Variable	N	Mean	Standard Deviation	Minimum	Maximum
Officially open	4189	0.25	0.44	0	1
Unofficial	4189	0.52	0.5	0	1
Experienced	4189	0.23	0.42	0	1
Distance	3733	3.05	1.91	0.27	9.06
HBcases	4189	2.76	8.75	0	93
Entries	4189	10.77	11.59	0	75
Foreclosures					
L0.	4189	27.84	45.37	0	1266
L3.	4012	28.08	45.95	0	1266
L9.	3658	25.85	42.48	0	1266
L12.	3481	24.8	41.14	0	1266
L15.	3304	23.97	40.08	0	1266
L18.	3127	22.87	38.77	0	1266

Table 2
Census Tract-Months and Census Tracts by HomeBase Program Eligibility
By Volume of Families Entering Shelter Between 1/2003 to 10/2004 and Time Period
Monthly Counts are Based on Observation Period between November 2004 and November 2008*

			Number of Families Entering Shelters from 1/2003 -- 10/2004					
	All Census Tracts		0,1,2		3 to 18		19 to 67	
	Is Census Tract Officially Eligible for Home Base Services?							
	Yes	No	Yes	No	Yes	No	Yes	No
Entire Study Period of HB Operations								
Census Tract Months: 11/2004 -- 11/2008	35,117	57,444	15,603	33,252	14,772	19,626	4,842	4,566
Census Tracts	*1,889*		*995*		*702*		*192*	
SubPeriods: Periods of Restricted Services and Citywide Expansion								
Restricted to Big 6 CDs: 11/2004 -- 6/2007	7,456	52,922	1,120	30,720	4416	18,048	1,920	4,224
Census Tracts	*233*	*1,656*	*35*	*96*	*138*	*564*	*60*	*132*
Expansion to all CDs: 7/2007 -- 11/2008	27,661	4,452	14,383	2,532	10,356	1,578	2,922	342
Census Tracts	*1,889*		*995*		*702*		*192*	
Start of HB services(Monthly counts are for 11/2004 to 11/2008 period)								
11/2004: CT Months	11,417	0	1,715	0	6,762	0	2,940	0
Census Tracts	*233*		*35*		*138*		*60*	
7/2007: CT Months	15,638	29,248	9,146	17,216	5,117	9,632	1,275	2,400
Census Tracts	*914*		*538*		*301*		*75*	
1/2008: CT Months	8,162	28,196	4,642	16,036	2,893	9,994	627	2,166
Census Tracts	*743*		*422*		*263*		*57*	

*For each census tract there were 22 months of observations on family entries into shelter, prior to the start of HB operations in November 2004.

Table 3

Number of Families Entering Shelter, Number of HB Cases Opened, and Average Number of HB cases/CT-Month by HB Eligibility, Volume of Families Entering System, Fiscal Year and Start of HB Eligibility

November 2004 to November 2008

	All Census Tracts		Number of Families Entering Shelters from 1/2003 -- 10/2004: 0,1,2		3 to 18		19 to 67	
	Is Census Tract Officially Eligible for Home Base Services?							
	Yes	No	Yes	No	Yes	No	Yes	No
Entire Study Period of HB Operations								
11/2004 -- 11/2008								
Shelter Entries	16,800	15,975	1,254	1,832	7,751	7,670	7,795	6,473
HB Cases	10,314	631	616	144	5,070	227	4,628	260
HB Cases/CT-month	0.29	0.011	0.040	0.0054	0.343	0.012	.956	0.057
SubPeriods: Periods of Restricted Services and Citywide Expansion								
Restricted to Big 6 CDs: 11/2004 -- 6/2007								
Shelter Entries	4,345	14,387	101	1,646	1,937	6,860	2,307	5,881
HB Cases	6,370	484	174	123	3,097	155	3,099	206
HB Cases/CT-month	0.854	0.009	0..156	0.004	0.701	0.009	1.614	0.049
Expansion to all CDs: 7/2007 -- 11/2008								
Shelter Entries	12,455	1,588	1,153	186	5,814	810	5,488	592
HB Cases	3,944	147	442	21	1,973	72	1,529	54
HB Cases/CT-month	0.143	0.033	0.031	0.008	0.191	0.046	0.523	0.158
Start of HB services								
11/2004								
Shelter Entries	7,770		201		3,439		4,130	
HB Cases	7,585		217		3,788		3,580	
HB Cases/CT-month	0.664		0.127		0.560		1.218	
7/2007								
Shelter Entries	5,441	7,625	660	881	2,476	3,523	2,305	3,221
HB Cases	1,681	342	269	96	705	108	707	138
HB Cases/CT-month	0.11	0.012	0.029	0.006	0.138	0.011	0.555	0.058
1/2008								
Shelter Entries	3,589	8,350	393	951	1,836	4,147	1,360	3,252
HB Cases	1,048	289	130	48	577	119	341	122
HB Cases/CT-month	0.128	0.010	0.028	0.003	0.199	0.012	0.544	0.056

Table 4
Effect of HB Capacity on Shelter Entries: CD Level Results
Linear specification, OLS regressions

HB capacity variables		
Officially operating H_{ct}	-2.36**	-2.48**
	(-5.89)	(-4.13)
Unofficially operating P_{ct}	-2.08**	-2.08**
	(-6.86)	(-6.84)
Experienced R_{ct}	---	0.16
	---	-0.27
Sum of HB capacity coefficients	-4.44**	-4.39**
	(-9.29)	(-8.66)
Sum of foreclosure coefficients	.053**	..053**
	(8.06)	(8.07)
Within R^2	0.34	0.34
N	*3127*	*3127*

t-values in parentheses. All regressions have month and CD fixed effects.
*Significant at 5 percent level.
**Significant at 1 percent level.

Table 5
Effect of HB Capacity on Shelter Entries: CD Level Results
Logarithmic specification, OLS regressions

HB capacity variables		
Officially operating H_{ct}	-0.066†	-0.06
	(-1.86)	(-1.16)
Unofficially operating P_{ct}	-.049†	-.049†
	(-1.88)	(-1.88)
Experienced R_{ct}	---	-0.009
	---	(-.018)
Sum of HB capacity coefficients	-.116**	-.118**
	(-2.80)	(-2.70)
Sum of foreclosure coefficients	.001*	.001*
	(2.52)	(2.51)
Within R^2	0.22	0.22
N	*3127*	*3127*

t-values in parentheses. All regressions have month and CD fixed effects.
**Significant at 1 percent level.
*Significant at 5 percent level.
† Significant at 10 percent level.

Table 6
Effect of HB Capacity on Shelter Entries: Stratified CD Level Results
Linear specification, OLS regressions

	Large CDs	Small CDs
HB capacity variables		
Officially operating H_{ct}	-2.10†	-0.168
	(-1.80)	(-0.39)
Unofficially operating P_{ct}	-0.93	-0.305
	(-1.13)	(-1.59)
Experienced R_{ct}	0.62	-0.169
	-0.53	(-.39)
Sum of HB capacity coefficients	-2.42*	-0.642
	(-2.26)	(-1.58)
Sum of foreclosure coefficients	.027**	.037**
	(2.67)	(5.19)
Within R^2	0.59	0.21
N	*1113*	*2014*

t-values in parentheses. All regressions have month and CD fixed effects.

**Significant at 1 percent level.

*Significant at 5 percent level.

† Significant at 10 percent level.

Table 7

Change in Rate of Family Shelter Entries Following Start of HB program: CT Level

HB Program Effects are estimated using fixed effects Poisson regression model for monthly observations between January 2003 and November 2008.

	Official Operations	Full Operations Model	Estimated reduction(-) or increase in Family shelter entries in HB eligible CTs/HB Cases Opened
	Incidence Rate Ratios (95% CI)		
All CTs			
Official Start of HB Operations	. .967	0.947	-573/10,314
	(.931,1.005)	(.881,1.018)	
Unofficial Start of Operations		0.973	
	.	(.929,1.010)	
3+ month Official Operations		1.03	
		(.958,1.106)	
Combined effect after 3+ months operations		0.949	-1,244/10,945
		(.897,1.005)	
CTs Grouped by Shelter Demand			
Low Demand CTs	1.13	1.25	514/760
	(1.04,1.22)	(1.14,1.38)	
Moderate Demand CTs	0.966	0.927	-995/5,297
	(..923,1.010)	(.871,.986)	
High Demand CTs	0.945	0.904	-1,297/4,888
	(..904,.989)	(.848,.964)	
Total Stratified Model			-1,778/10,945
CTs Grouped by Official Start of HB Operations			
11/2004	0.996	0.97	
	(.946,1.049)	(.91,1.04)	
7/2007	0.914	0.89	
	(.870,.959)	(.83,.95)	
1/2008	1.01	1.01	
	(.960,1.071)	(.94,1.09)	

Source:hb12poisson11_23)

Table 8

Change in Monthly Family Shelter Entries Following Start of HB program: CT Level

HB Program Effects are estimated using fixed effects linear regression model for monthly observations between January 2003 and November 2008.

	Official Operations	Full Operations Model	Estimated reduction(-) or increase in Family shelter entries receiving HB services
	Linear Effects (95% CI)		
All CTs			
Official Start of HB Operations	-0.07144	-0.056	-2,493
	(-.089,-.052)	(-.071,-.007)	
Unofficial Start of Operations		-0.071	
	.	(-.085,-.056)	
3+ month Official Operations		-0.003	
		(-.030,.025)	
Combined effect after 3+ months operations		-0.130	-7,184
		(-.153,-.106)	
18-month foreclosure	0.041	0.038	
	(.033,.049)	(.030,.046)	
CTs Grouped by Shelter Demand			
Low Demand CTs	-0.01	-.01	-20
	(-.021,.000)	(-.025,.006)	
Moderate Demand CTs	-0.017	-0.051	-1,240
	(-.043,.009)	(.-.093,-.01)	
High Demand CTs	-0.003	-0.114	-1,013
	(-.086,081)	(-.270,.043)	
Total Across Strata			-2,273
CTs Grouped by Official Start of HB Operations			
11/2004	0.04	-.03	
	(.018,.062)	(-.08,.02)	
7/2007	-0.11	-0.16	
	(.-.129,.091)	(-.18,-.13)	
1/2008	-0.069	-0.1	
	(-.092,-.046)	(-.15,-.08)	

Source: hb12c (December 29, 2010)

Table 9
Estimates of the Historical Effect of HomeBase on Shelter Entries
Different Estimates of HB Effect per 100 HB Cases

	CD			CT		
	Point	Bottom	Top	Point	Bottom	Top
Capacity						
Linear	-60.5	-73.2	-47.8	-65.4	-76.2	-54.6
Stratified model	-19.2	-34.2	-4.3	-20.8	-34.4	-7.2
Log/poisson	-25.1	-46.8	-3.5	-11,8	-25.4	1.7
Service						
Linear	-9.9	-13.9	-5.9	-12.0	-15.0	-9.0
Quadratic	-26.1	-35.2	-16.9	--	--	--

Note: Point estimates are calculated by estimating the number of entries averted for each equation during unofficial and official HB operations adjusted for the experience factor. This total is divided by the total number of HB cases open and multiplying the ratio 100. Confidence intervals are calculated using the Stata lincom command.

Table 10
Effect of HB Service on Shelter Entries: CD Level Results
Instrumental Variable Regressions

HB families served	-.0989*	-.5240*
	(-4.89)	(-4.70)
(HB families served)^2	---	.00865*
	---	(3.88)
R^2	0.86	0.85
N	*2725*	*2725*

t-values in parentheses. All regressions have month and CD fixed effects, and lagged foreclosures.

*Significant at 1 percent level.

Table 11

Families Entering Shelters from NYC Census Tracts between 11/1/2004 and 11/30/2008: CT Level Service Model

Instrumental variable, fixed effect linear model

All CTs				
Monthly HB cases opened	-0.12			
	(-0.15,-0.09)			
6-month cumulative effect of opening an HB case		-0.30		
		(-0.49,-o.18)		
Monthly HB Cases Stratified by Shelter Demand				
Low Demand			-0.15	
			(-0.30,0.00)	
Moderate Demand			-0.029	
			(-.082,.025)	
High Demand			-0.039	
			(-0.10,0.022)	
Monthly HB Cases Stratified by Official Start Date				
11/2004				-0.09
				(--0.13,-.056)
7/2007				-0.24
				(-0.36,-0.12)
1/2008				0.20
				(.085,0.32)
18-month Cumulative Effect of Additional Unit Entering Foreclosure	0.046	0.06		0.045
	(0.038, 0.054)	(0.052, 0.073)		(0.036,0.054)

Source: myhb08d, May 23, 2011.

.

*p<.0.5,**p<.01,**p<.001

Table 12

Effect of HB Estimated with Larger Units

Linear specification

	Single month
Capacity (OLS)	
Single CD	-4.44*
	(-9.29)
Double CD	-5.34**
	(-9.17)

t-values in parentheses. All regressions have month and CD fixed effects and foreclosures. The capacity effect is the sum of the coefficients on unofficial and official operation.

**Significant at 1 percent level.

Table 13
HB Capacity and Service Equation Effects
for 1-, 2-, 3-, and 6-month Grouping of Observations

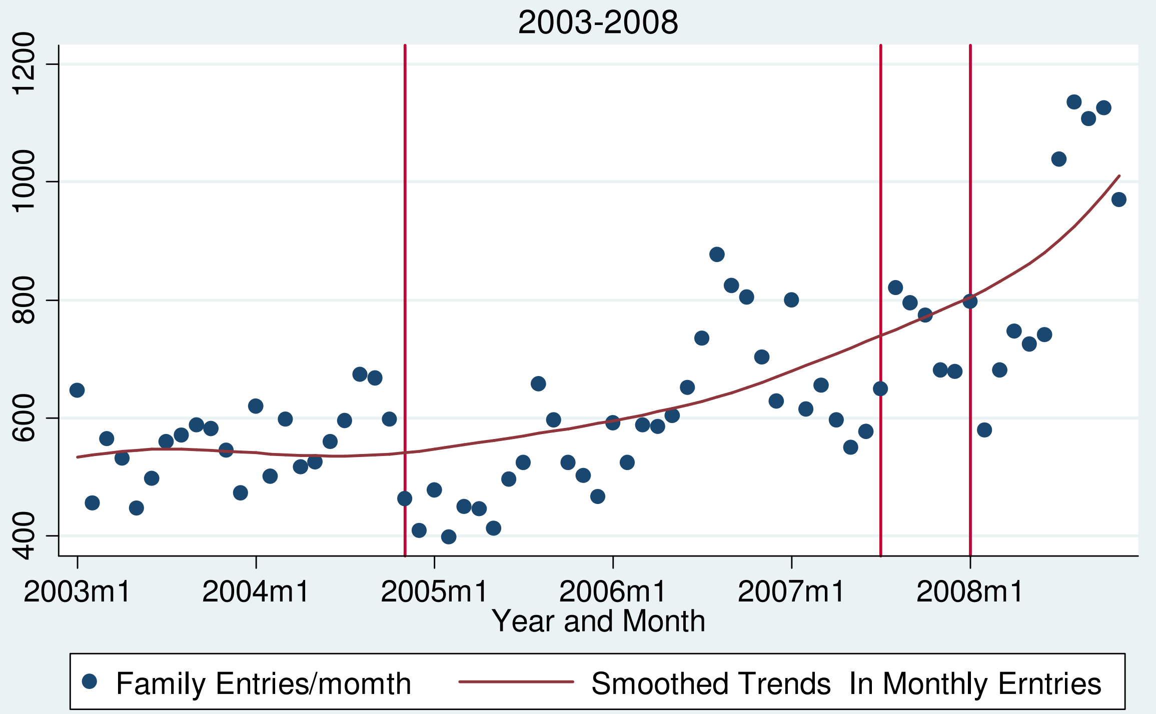
	One month	Two month	Three month	Six month
Capacity Equation Poisson	0.952	0.952	0.954	0.949
	(.899,1.008)	(.898,1.010)	(.899,1.012)	(.891,1.012)
Services equation	-0.21	-0.18	-0.19	-0.18
	(-.24,-.17)	(-.21,-.14)	(-.23,-.15)	(-.21,-.14)

Table 14
Effects of HomeBase Capacity on Exit Hazards
Cox Proportional Hazard Results

	Non-placement effect β	$\tilde{\beta}$	Placement effect $(\beta + \tilde{\beta})$
Official operation	0.024	-0.056	-0.032
	(0.52)	(-0.76)	
Unofficial operation	-0.004	0.019	0.015
	(-0.08)	(0.38)	
Experienced	0.003	0.018	0.021
	(0.08)	(0.25)	
Sum	0.023	-0.019	0.004
Log pseudolikelihood			-426,316.72
N			*90,222*

z-statistics in parentheses.

Figure 1: Trends in Family Entries in the NYC Shelter System



Vertical Lines mark the phased start of offical HB operations

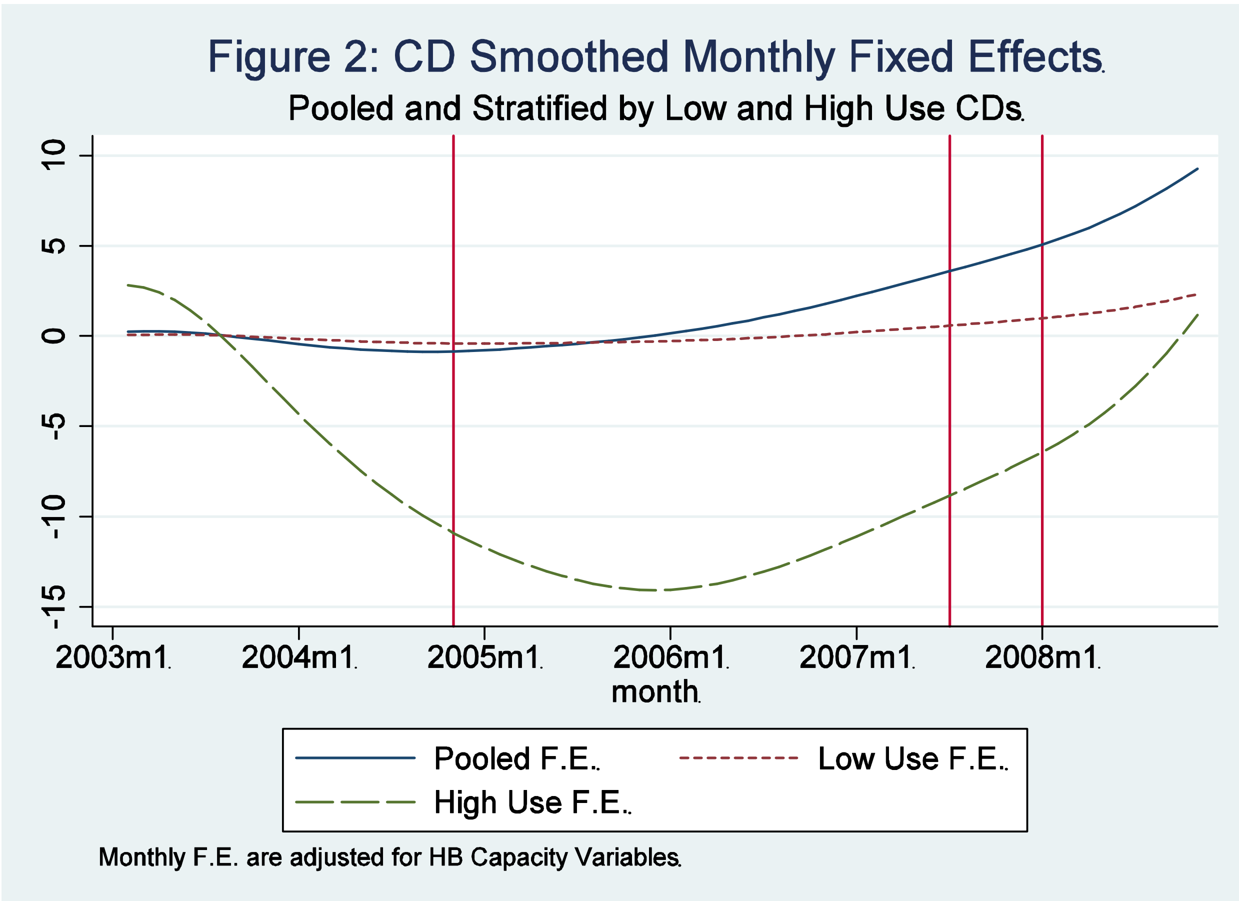


Figure 2: CD Smoothed Monthly Fixed Effects
Pooled and Stratified by Low and High Use CDs
10
5
0
-5
-10
-15
2003m1
2004m1
2005m1
2006m1
2007m1
2008m1
month
Pooled F.E.
Low Use F.E.
High Use F.E.
Monthly F.E. are adjusted for HB Capacity Variables

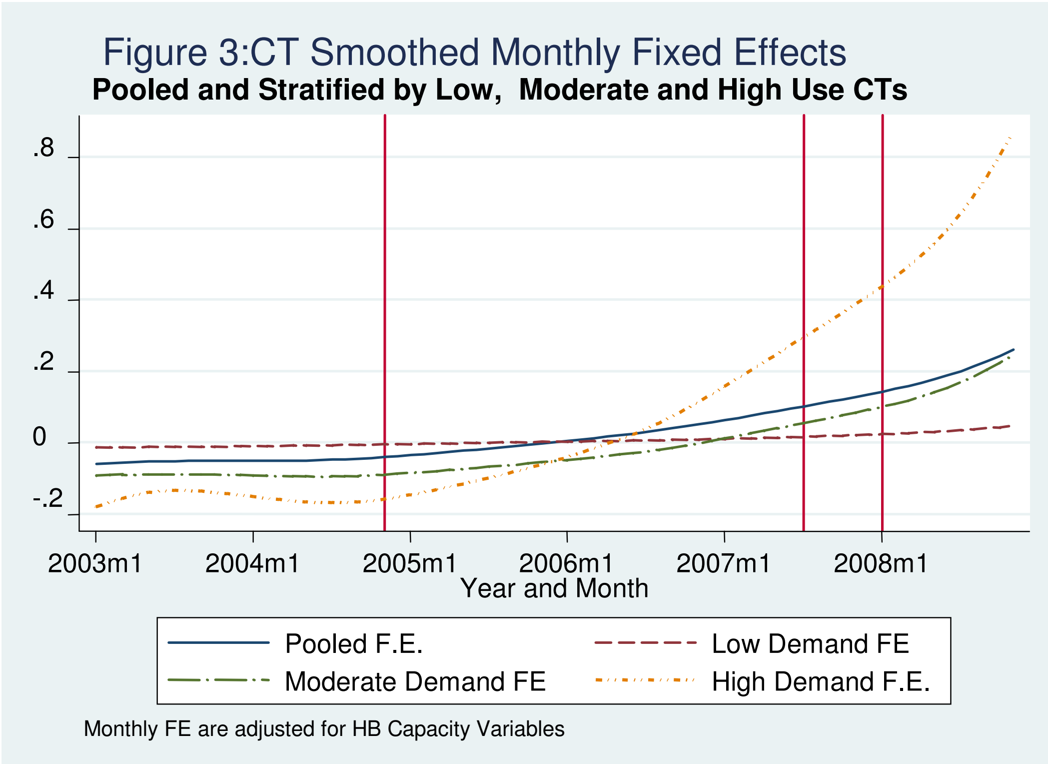


Figure 3:CT Smoothed Monthly Fixed Effects
Pooled and Stratified by Low, Moderate and High Use CTs
.8
.6
.4
.2
0
-.2
2003m1
2004m1
2005m1
2006m1
2007m1
2008m1
Year and Month
Pooled F.E.
Low Demand FE
Moderate Demand FE
High Demand F.E.
Monthly FE are adjusted for HB Capacity Variables

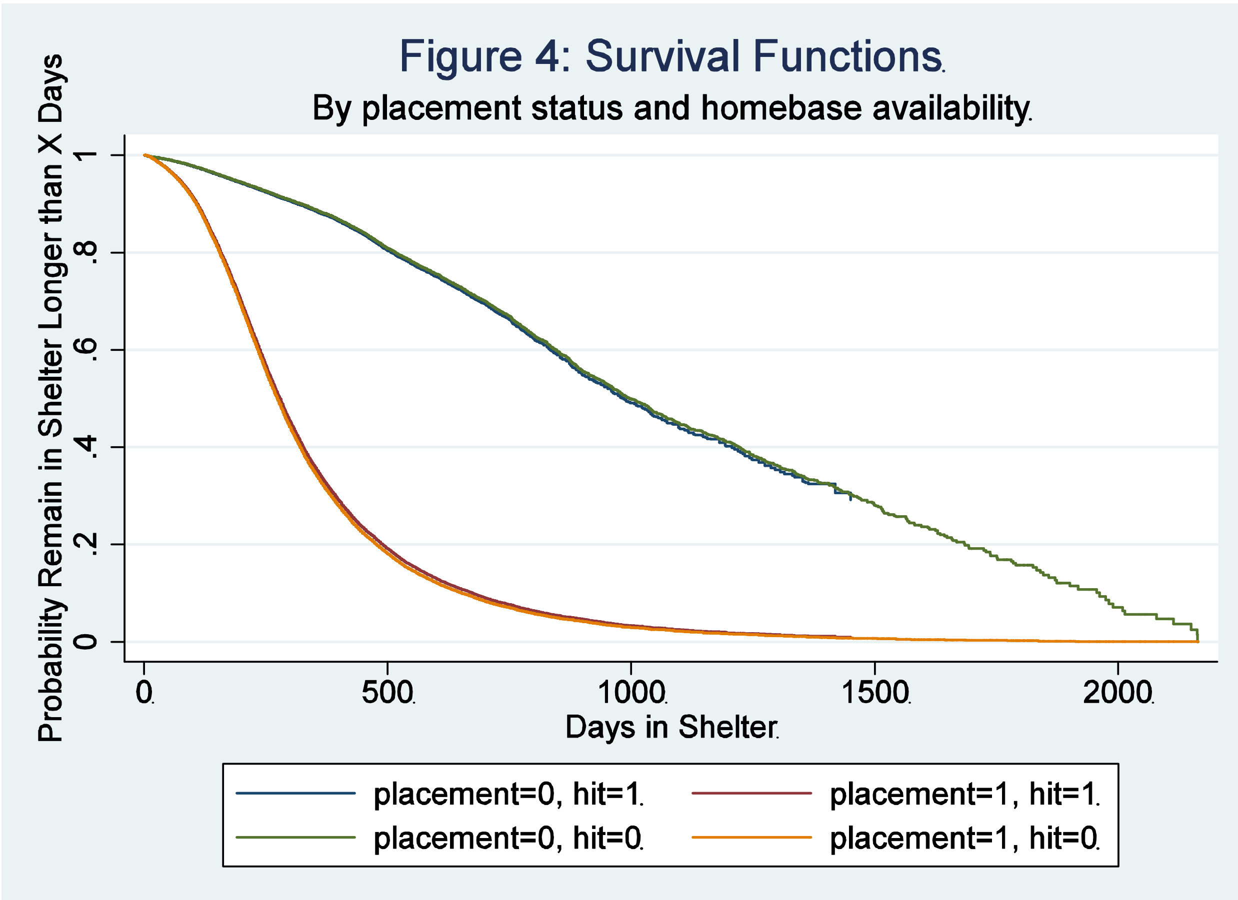


Figure 4: Survival Functions
By placement status and homebase availability
Probability Remain in Shelter Longer than X Days
0
.2
.4
.6
.8
1
0
500
1000
1500
2000
Days in Shelter
placement=0, hit=1
placement=1, hit=1
placement=0, hit=0
placement=1, hit=0