

Aggregate-Level Inferences from Individual-Level Data:

The Case of Permanent Supportive Housing and Housing First

Brendan O’Flaherty

Department of Economics

Columbia University in the City of New York

This draft: January 15, 2022

Abstract

I estimate the “simple mechanical effect” of permanent supportive housing and Housing First as studied in the At Home/Chez Soi and HUD-VASH experiments on point-in-time counts of homelessness (HUD definition). The simple mechanical effect is the effect that would occur in the absence of any behavioral responses aside from those in the experiments. The estimates of the simple mechanical effects overlap the confidence intervals in Corinth’s (2017) regression study of the total effect. This finding suggests that the net effect of behavioral responses outside the experiments is small. The essay illustrates how useful inferences about aggregate-level phenomena can be derived from individual-level data.

I am grateful to Emmy Tiderington and Yi-Ping Tseng for helpful comments and information.

Strategies that work on an individual level do not always work the same way on an aggregate level. Running your air conditioner on high all summer may be a good way to keep yourself cool in the era of climate change, but it will not keep the planet cool (given today's electricity generating technology). Stockpiling groceries and toilet paper at the start of a pandemic may be a wise precaution for a household to take, but if everyone runs to the supermarket to stockpile, the pandemic may spread faster and the rise in demand may create a shortage that would not otherwise have existed; stockpiling may create and exacerbate the problems it was trying to avoid.

Similarly, interventions that produce promising results on an individual basis do not translate automatically into policies that reproduce the same results on an aggregate level. Translation is difficult and uncertain. This essay is about an example of such translation. It shows how translation can be done, what its uncertainties are, and how it can be valuable.

The effects of permanent supportive housing in general and (Pathways) Housing First¹ in particular are areas where the implications of individual level patterns on aggregate level outcomes have not always been clear, and have sometimes been stated incorrectly. Several high-quality randomized controlled trials (RCTs) have shown that permanent supportive housing and Housing First cause improvements in many housing outcomes for their participants (see reviews in, for instance, Kertesz and Johnson 2017 and Miler et al. 2021), but these are individual-level results. Aggregate-level results would require RCTs that randomly assigned permanent supportive housing (PSH) and Housing First (HF) beds to different cities, or natural experiments where such beds were

¹ In this essay when I say "Housing First" or "HF," I mean Pathways Housing First.

assigned as if randomly. So far, no such RCT has been reported, and only one viable natural experiment has been studied (Corinth 2017).²

Hence how much (if at all) permanent supportive housing or Housing First reduce aggregate homelessness is not well known at this point. But some of the RCTs have data that allow us to make some inferences about the size of the effects they might have. These RCTs allow us to estimate what I will call the “simple mechanical effect” of these interventions on point-in-time (PIT) counts of homelessness (in this case, homelessness according to the U.S. definition). The simple mechanical effect means the effect the intervention would have if the only behavioral responses were those that the RCT studied. The simple mechanical effect is neither an upper bound nor a lower bound on the complete effect: the behavioral response outside the RCT might reduce homelessness (contagion effects, for instance) or raise it (incentives to remain homeless, for instance). Instead, the simple mechanical effect is a starting point for analysis if the net outside-RCT behavioral effects are large, or an approximation to the complete effect if the outside-RCT behavioral effects are small.

Looking at two RCTs in depth

Calculating the simple mechanical effect begins with information in the RCT about the causal effect of the intervention on the proportion of days homeless. Relative to any participant’s entrance into the RCT, the day of a PIT count is essentially random, and so the expected effect of participation in a particular stage of the RCT is estimated by the difference between treatment and control groups at that stage in the proportion of days homeless.

² Evans et al. (2019) use a shift-share instrument involving veteran population across continua-of-care to study the effect of award of HUD-VASH vouchers, but this instrument violates the exclusion criterion. Funding from several other programs to combat veteran homelessness, especially Social Services for Veteran Families, is likely to be positively correlated with this instrument as well.

Two respected RCTs report this information: At Home/Chez Soi (Goering et al. 2014) and the HUD-VASH experiment (Rosenheck et al. 2003), and I will concentrate on these two studies. In neither case was this difference a primary concern of the RCT, and some RCTs do not report this difference at all. (Goering et al. in fact report days in emergency shelters and days on the street, and I sum these two categories to create a measure of days homeless under the US definition. Canadians do not have to use US definitions.)

In At Home/Chez Soi, the primary outcome of interest is days “stably housed” and in the HUD-VASH experiment the primary outcome of interest is days housed. In neither case is the proportion of days homeless merely the difference between one and the proportion of days in the primary outcome. In the HUD-VASH experiment the third category is days institutionalized, and in At Home/Chez Soi there is a fourth category, days in “temporary housing.” “Stable housing” is defined as “living in one’s own apartment or house, with family, with an expected duration of residence greater than or equal to six months and/or tenancy rights” (p. 43). Housing that is not stable is temporary, and so includes at least couch-surfing, SROs, rooming houses, hotels, and motels.

The distinction between days homeless and days not (stably) housed is quite important: both interventions cause considerably larger decreases in days not (stably) housed than in days homeless. Table 1 shows how the categories compare in the two RCTs.

(Table 1 around here.)

The most notable feature of table 1 is that the reduction in homelessness is considerably less than the increase in (stable) housing: 15% versus 41% in At Home/Chez Soi; 8.2% versus 13.1% in the HUD-VASH experiment. Both interventions were successful in cutting homelessness, but in

neither case did the control group spend a great deal of time homeless: 24% in At Home/Chez Soi and 22.9% in the HUD-VASH experiment.

(Table 1 also shows that the difference in outcomes was not as large as the headline numbers suggest: At Home/Chez Soi raising stable housing by 41% vs HUD-VASH raising housing by 13.1%. Counting outcomes in At Home/Chez Soi the same way they were counted in HUD-VASH reduces the 41% effect to 20%.)

Thus, a first, rough estimate of the simple mechanical effect of adding 100 participants to permanent supportive housing or Housing First and reducing treatment as usual by 100 would reduce the PIT count by 8 to 15. Even the most efficacious intervention imaginable with these participants would reduce homelessness by 23 or 24 per hundred—not by a hundred.³ It is clear from the design and reporting on these experiments that reducing the PIT count was not the primary goal of these interventions. Nothing in these RCTs indicate that permanent supportive housing or Housing First is likely to cause large reductions in the PIT counts, although modest reductions are supported.

From RCTs to actual programs

The simple averages in table 1 cannot adequately estimate the simple mechanical effect on the PIT count because the effects on homelessness are not invariant to time in the program, and because programs in actual operation will not have the same distribution of participants by time in the program that the RCTs used to find the numbers in table 1.

Two major differences between the RCTs and actual programs work in opposite directions. On one hand, actual programs will have participants who entered the program more than two years

³ These numbers are remarkably close to those in Rolston et al. (2013), an RCT of the HomeBase homelessness prevention program in New York City.

ago (At Home/Chez Soi) or three years ago (HUD-VASH). As time in program increases, the effect on homelessness is likely to decrease, largely because control participants leave homelessness, either positively (finding a home) or negatively (becoming institutionalized or dying); this would make the treatment effect on homelessness smaller. On the other hand, actual programs usually respond to program attrition by recruiting new participants, and so a program with turnover may have treatment participants with short tenure than an RCT that follows the same people whether they are currently enrolled in the program or not.

To sort out these possibly opposing effects, we need a model of how a permanent supportive housing or Housing First program would actually operate, and how that operation would be reflected in PIT counts. For this, we need three components: an account of how the program effect on days homeless changes with time-from-entry into the program (the “program-homelessness” effect), an account of how participants leave the program (the “program-attrition-rate”), and an account of how program capacity changes.

Start with the program-homeless effect, and denote time periods by $\sigma = 0, 1, 2, \dots$. Let $\Delta(\sigma) \in (-1, 1)$ denote the difference in the rate of homelessness (proportion of time) between treatment and control participants at σ periods from starting the program (whether or not the treatment participant is still in the program). We define $\Delta(\sigma)$ as *positive* if the program reduces the proportion of days homeless. This is what the RCTs measure.

For an account of how participants leave the program, we will assume a constant per period hazard of leaving of $\delta \geq 0$, called the program attrition rate. (The program attrition rate is not the rate of attrition from an RCT; the program-homeless effects from the RCTs include participants who have left the program as long as the RCT can interview them.) A constant hazard rate is a strong assumption, but it greatly simplifies the analysis. We need to model leaving the program

because it tells us how many new participants can enter each period without changing the program capacity. If we used a time-varying hazard rate, we would have to produce an experience profile of the program population every period. That exercise would distract from the main message of this paper.

For an account of program capacity, we assume that in each period t there is an increase $C_t \geq 0$ in capacity. (We do not worry about modelling decreases in capacity, but accommodate long sequences of periods with no change in capacity.) We assume that at the beginning of every period, new participants are brought in to eliminate vacancies.

To see how these elements fit together, suppose there is an increase in capacity of C_0 in period 0, and no further changes. Every period after that, then δC_0 participants leave and δC_0 new participants are brought on board. At time $s > 0$, then, there are C_0 participants who entered at time 0; these participants (whether they are in the program or not) cause a decrease $C_0\Delta(s)$ in the PIT count of period s . In addition, there are δC_0 participants who entered in period 1; they cause a reduction of $\delta C_0\Delta(s - 1)$ in the PIT count. There are in fact δC_0 participants who entered in every period up to and including s , and they also cause decreases in PIT homelessness. Putting the participants from different cohorts together, the decrease in PIT homelessness at time $s > 0$ is

$$C_0\Delta(s) + \sum_{\sigma=0}^{s-1} \delta C_0\Delta(\sigma) = C_0\{\Delta(s) + \delta \sum_{\sigma=0}^{s-1} \Delta(\sigma)\}.$$

At time $s = 0$, we assume that new participants enter only to fill the new beds, not as replacements, and so the decrease in PIT count is $C_0\Delta(0)$.

These expressions easily generalize to multiple expansions of program capacity. Each expansion generates its own set of new participants: C_t in period t , and then δC_t in every

subsequent period. Let $D(s)$ denote the decrease in PIT count that the program causes in period s .

Then we can write

$$(1) \quad D(s) = C_s \Delta(0) + \sum_{t=0}^{s-1} C_t \{\Delta(s-t) + \delta \sum_{\sigma=0}^{s-t-1} \Delta(\sigma)\}$$

We will use equation (1) to simulate the various plausible scenarios on a sequences of PIT counts.

Notice that (1) is linear in the sequence of capacity expansions (C_t). If all capacity expansions change by the same proportion, all reductions in PIT count change by that same proportion.

To understand how equation (1) works, it is useful to take a first difference:

$$(2) \quad D(s+1) - D(s) = C_{s+1} \Delta(0) + C_s [\Delta(1) - \Delta(0)] + \sum_{t=0}^{s-1} C_t \{\Delta(s-t+1) - (1-\delta)\Delta(s-t)\}.$$

Since each C_t is a first difference in capacity, this equation implies that first differences in PIT count should not be regressed on the first difference in capacity, as is normally done, but on all previous first differences in capacity.

For $C_{s+1} = C_s = 0$, the first difference in PIT count simplifies to

$$D(s+1) - D(s) = \sum_{t=0}^{s-1} C_t \{\Delta(s-t+1) - (1-\delta)\Delta(s-t)\}.$$

Note that the expression in curly brackets is positive if and only if

$$\delta > \frac{\Delta(s-t) - \Delta(s-t+1)}{\Delta(s-t)}$$

That is, if and only if the rate of program attrition is greater than the rate at which the program-homelessness rate decays. The expression is negative when the decay rate is greater than the

attrition rate. Hence the overall effect on PIT count may either increase or decrease over time
s. Decreases are less common when program attrition rate is greater.

Calibration

To calibrate equation (1), we need plausible values for the program-homelessness effects, the program-attrition rate, and the sequence of program expansions. We also need a unit of time.

For the program-homelessness effects, we rely on the two RCTs, At Home/Chez Soi and the HUD-VASH experiment. Because At Home/Chez Soi reports quarterly effects, we will use quarters as our unit of time. Both RCTs report these effects on figures, rather than tables (figure 4, page 19 in Goering et al., and figure 2, page 946 in Rosenheck et al.). I approximated these values from the figures. At Home/Chez Soi had two study groups: a high needs group who received Assertive Community Treatment (ACT), and a moderate needs group who received Intensive Case Management (ICM). I treat each study group separately. The program-homeless effects were considerably greater for the ICM group than for the ACT group (whose effects were similar to those of HUD-VASH), partly because the ACT control group spent a great deal of time in institutions and little time homeless.

Three difficulties were encountered in this process.

First, Goering et al. publish no data on the time sequence of program-homelessness effects on street-living (probably because the numbers are quite small). I assumed that this effect was constant over time and across study groups at 5% (the average value from table 1).

Second, the HUD-VASH experiment reports at 6-month and 12-month intervals at some times, not quarterly. I filled in blanks with the next value reported.

Finally, At Home/Chez Soi reports on only two years, and the HUD-VASH experiment, only three years. I have no information on long run program-homelessness effects.

Two papers, however, report on *housing stability* results for six years for Housing First programs. Kuehnle et al. (2021) studied a program in Melbourne, Australia (Journeys to Social Inclusion, or J2SI) and found that after six years there was no difference between treatment and control groups in the proportion “securely housed,” even though the difference was about the same as that of At Home/Chez Soi after two years. On the other hand, Stergiopoulos et al., in following the Toronto participants from At Home/Chez Soi, found that the treatment effect on housing stability fell over time but did not disappear after six years. For the high-needs ACT group, which had a small decrease in homelessness, the difference in days stably housed fell from about 38% in year two to about 25% in year 6 (table 2, page 922). For the moderate-needs ICM group, which had a larger difference in homelessness, the difference in days stably housed fell from about 37% in year two to about 10% in year 6, a difference that does not appear to be statistically significant (table 3, page 922). In all cases, the differences decrease because a greater proportion of the control group becomes stably housed as time passes.

For the At Home/Chez Soi study groups between 24 and 36 months, I assigned an effect equal to half of the effect at 24 months, because that was the pattern in the HUD-VASH experiment. For all groups between 36 and 72 months, I kept the effect the same effect the same as it was at 36 months (which was 5.025% for the ACT group, 7.55% for the ICM group, and 4% for HUD-VASH). I assigned a zero effect for everything beyond six years (72 months).

For the program attrition rate, my main sources are Wong et al. (2006), which looks at permanent supportive housing for people with serious mental illness around the turn of the century, and McClure (2017), which looks at all HUD-assisted housing from 1995 to 2015. The two sources

present information in different ways, but neither is informative about attrition after more than a decade. For comparability, I translated each to a leaving rate per month for different periods. In the HUD-assisted housing data, I concentrated on non-elderly housing choice voucher (HCV) recipients without children. McClure found that income had no effect on length of stay in HUD-assisted housing generally, and so these rates might be meaningful for formerly homeless people. I looked both at spells that began around 1995 and spells that began around 2010. McClure noted that spells tended to become longer for all forms of assisted housing over this time period.

Table 2 compares the implied monthly program-attrition rates for these three groups: Philadelphia permanent supportive housing residents with spells beginning around 2000, non-elderly HCV recipients without children with spells beginning around 1995, and non-elderly HCV recipients with spells beginning around 2010.

(Table 2 around here.)

All of the rates are between 0.98% and 2.82%, and later rates are generally lower than earlier. The permanent supportive housing rates at the turn of the century are mainly lower than the HCV rates then, and so permanent supportive housing rates later than the turn of the century are probably lower than HCV rates later in the century, which are generally quite low. For the simulations, I used constant monthly attrition rates of 1.5% and 2.0%.⁴

Finally, to calibrate the sequence of capacity expansions, I use two alternative historical sequences: the new HUD-VASH vouchers awarded to public housing authorities from fiscal year 2008 through fiscal year 2020 (U.S. Department of Housing and Urban Development, 2021a); and

⁴ J2SI had a program attrition rate of 20% over three years, or about 0.6% per month (Kuehnle et al 2021). This is an actual Housing First program, albeit small. Hence the proposed program-attrition rates may be on the high side.

the increase in the number of permanent supportive housing (year-round) beds reported in HUD's Housing Inventory Count (U.S. Department of Housing and Urban Development, 2021b) from 2007 to 2020. In this time period, HUD awarded 104,966 new housing choice vouchers under HUD-VASH, and the permanent supportive housing bed count rose from 188,636 to 373,030.

Figure 1 shows the two capacity expansion sequences. Although beds supported by HUD-VASH vouchers are a subset of permanent supportive housing beds in the Housing Inventory Count, the two sources are conceptually different and have different timing conventions; they are not directly comparable. Both sequences show a steady expansion with peaks around 2011 and 2016, and a much slower expansion after 2016. The permanent supportive housing expansion is much more volatile than the HUD-VASH expansion, and its decline after 2016 is greater.

Thus we will be simulating the change in PIT counts caused by the two capacity expansions. Because of the linearity of equation (1) in the sequence of capacity expansions, we can interpret the resulting patterns as those of a representative continuum of care. Replicating the aggregate historical change in PIT counts is not the main goal. The main goal is to see how PIT counts respond (in a simple mechanical way) to sorts of capacity changes a representative continuum of care may have experienced. We are more interested in relative than absolute changes.

Results

I performed six different simulations. I always combined At Home/Chez Soi program-homelessness effects with the PSH capacity expansion, and HUD-VASH program-homelessness effects with the HUD-VASH program expansion. For each of the resulting three combinations, I simulated once with a monthly program attrition rate of 1.5% and once with 2.0%. So I name the simulations with a combination of a study group (ACT, ICM, or HUD-VASH) and a monthly program attrition rate (1.5% or 2%).

The resulting implied changes in PIT homelessness are shown in figures 2 (ICM), 3 (ACT) and 4 (HUD-VASH). These figures, then, are the simple mechanical effect of permanent supportive housing and Housing First on PIT homelessness (using the HUD definition).

(Figures 2, 3, and 4 around here.)

In all three figures, the higher program attrition rate is associated with greater reductions in PIT homelessness, with the gap growing over calendar time. The reason is obvious: higher program attrition means that a program sees more people, holding capacity constant, and seeing more people results in more opportunities to reduce homelessness. This result, however, should not be interpreted to imply that programs with higher attrition are better at reducing homelessness. The result is mainly an artifact of how I have pieced the simulation together, with the search for the appropriate program attrition rate separate from the search for the appropriate program-homelessness rate. In reality, the two rates are unlikely to be independent: one probably cannot manipulate program attrition rates without simultaneously affecting the ability of the program to avert current and future homelessness. Various Moving On initiatives may increase the program attrition rate without increasing homelessness, but they almost always cost money, frequently in the form of Housing Choice Vouchers (Tiderington et al. 2022).

The jumps in averted homelessness every four quarters are also an artifact of how I constructed the simulations. I put the entire capacity expansion for each year in the first quarter. Smoothing the capacity expansion would smooth the graphs—but also reduce the effect of the capacity expansions.

Even though capacity grows monotonically under both scenarios the reduction in PIT count caused by that (cumulative) expansion does not grow monotonically in any simulation. As time passes, the PIT count reductions that older cohorts cause diminish, and can be offset only by new

capacity expansions. When the size of capacity expansions also diminishes in later years, as it does in both scenarios, the total effect gets smaller, even though the cumulative amount of capacity expansions keeps growing.

Because the permanent supportive housing scenario slows more than the HUD-VASH scenario does, the simulations associated with At Home/Chez Soi reach their peak in homelessness reduction between 8.75 and 12 years in the 13-year scenarios. With the HUD-VASH scenario, the global maximum is reached at the end, but local maxima occur before that.

Regressions in first differences

A naïve reaction to a sequence of capacities and a corresponding sequence of PIT count reductions is to regress first differences in one sequence on first differences in the other sequence to see how one sequence is impacting the other. Corinth (2017) and Evans et al. (2019) both follow essentially this strategy, although in much more sophisticated ways (because they are confronting actual data, not data they simulated). That is, in our notation,

$$(3) D(s + 1) - D(s) = \beta C(s + 1) + \varepsilon$$

where the object is to estimate the regression coefficient β . We know from equation (2) that equation (3) is mis-specified, but since (2) is extremely hard to estimate, especially with the short panels that are available now, the approach is understandable.

Consider a sequence of observations from period 0 through period T . Let $\hat{\beta}_T$ denote the ordinary least squares estimator of β on that data set. This takes a particularly simple form:

$$\hat{\beta}_T = \frac{D(T)}{\sum_{t=0}^T C_t}$$

the ratio of the current reduction in the PIT count to cumulative capacity expansions.

This ratio is of independent interest. The reduction in PIT count at time T , $D(T)$, has been produced by capacity expansions up to and including that time, and so it seems natural to use the ratio as a measure of the strength of the relationship.

Figure 5 therefore shows the ratio $\hat{\beta}_T$ over 13 years (at one-year intervals) for all six simulations. In all cases, the lines slope downwards, as expected from the concave downward shape of the simulation results: over time, cumulative capacity expansion rises more quickly than the PIT count reductions. The largest values of the ratio occur for the ICM 2% simulation, which starts at 0.248 and falls to 0.155. The smallest values are for the HUD-VASH 1.5% simulation, which starts at 0.172 and falls to 0.085. The vast majority of values are in the [0.10, 0.20] range.

(Figure 5 around here.)

These values are not inconsistent with Corinth's (2017) results from actual PIT and Housing Inventory Count data. He studied a seven-year period, 2007-2014, but since he worked with first differences, it would be considered a six-year period for my purposes. He summarizes his results by saying that "one additional PSH bed reduces homeless counts by up to 0.10 people," but his standard errors indicate 95% confidence intervals that include the entire [0, .20] range. After six years, my simulations for $\hat{\beta}_T$ vary from 0.202 for ICM 2.0 to 0.118 for HUD-VASH 1.5.⁵

Remember that these simulations are intended to estimate the simple mechanical effect of the capacity expansions, not the actual effect. If they do indeed estimate that effect, then Corinth's estimates of the actual effect suggest that outside-of-RCT behavioral responses are small on net.

⁵ The Evans et al. (2019) estimates are much bigger, around one. I noted in footnote 1 that these are really estimates for the whole panoply of veteran initiatives, not just HUD-VASH. They are also upwardly biased because they do not account for the aging of Vietnam-era veterans. See O'Flaherty (2019, 2022) for a more detailed discussion.

Of course, the simulations may not accurately reflect the simple mechanical effect. There could be problems with both external validity of the RCTs and with the calibrations that led from the RCTs to the changes in PITT counts. The external validity concerns are about actual program operation—how close were they to following Housing First protocols, for instance?⁶--and about the characteristics of participants—were they like the participants in the two RCTs?⁷ The calibration concerns arise because I have very little data beyond the first few years. But the estimates of the simple mechanical effect seem reasonably robust—values of $\hat{\beta}_T$ around 0.40 or over seem unlikely.

Conclusion

Individual-level results by themselves cannot support aggregate-level conclusions. That is obvious. But they can discipline those conclusions, and enrich our understanding of them. Goering et al. (2014) is solid science, and so is Corinth (2017), and we should (tentatively, of course) accept both of them, even though their elevator-speech summaries (“Housing First is good;” “Housing First is bad”) appear inconsistent. I have shown that they are not inconsistent, and each is a more valuable contribution because of the other.

Methodologically, this essay has shown how to go from individual-level results to aggregate-level conclusions, and what sort of parameters need to be estimated better to make the translation more accurate.

Substantively, I have shown that Housing First does not come close to reducing the PITT count one-for-one, but nothing else does either. In the real world, reductions of the PITT count like

⁶ Tiderington (2019) notes that the At Home/Chez Soi reports leave open the precise services the participants received, and probably more importantly, those that constituted “treatment as usual.”

⁷ Milburn et al. (2021), for instance, show that much PSH in Los Angeles deviates considerably from At Home/Chez Soi. The Los Angeles CoC accounted for 6.1% of all PSH beds in 2020.

those in the simulations or like those in Corinth's paper are extraordinarily good (for comparisons, see Ellen and O'Flaherty (2010, page 12, note 3) and Corinth (2017, page 79). Outside the reductions in PIT count, the reductions in unstable housing conditions are also valuable—and they would be counted as reductions in PIT count in the more expansive definitions of homelessness that are common outside the US.

Efficacious vaccines did not “end COVID,” either. I thought that they would not, but I still got my shots and booster as soon as I was eligible.

References

Corinth, Kevin, 2017, The impact of permanent supportive housing on homeless populations, *Journal of Housing Economics* 35: 69-84.

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

Evans, W N, Kroeger S, Caroline Palmer C, and Emily Pohl R (2019) Housing and urban development—Veterans Affairs supportive housing vouchers and veterans' homelessness, 2007-2017. *American Journal of Public Health* 109(10): 1440-1445.

Goering, Paula, Scott Veldhuizen, Aimee Watson, Carol Adair, Brianna Kopp, Eric Latimer, Geoff Nelson, Eric MacNaughton, David Streiner, and Tim Aubry, 2014, *National At Home/Chez Soi Final Report*. Calgary, AB: Mental Health Commission of Canada.

Kertesz, SG, Johnson G (2017) *Housing First: Lessons from the United States and Challenges for Australia*. *Australian Economic Review* 50(2), 220-228.

Kuehnlé, Daniel, Guy Johnson, and Yi-Ping Tseng, 2021, Making it home? Evidence on the long-run impact of an intensive support program for the chronically homeless on housing, employment, and health. Working paper, University of Duisburg-Essen & CINCH, Essen, Germany.

McClure, Kirk, 2017, Length of stay in assisted housing. U.S. Department of Housing and Urban Development, Office of Policy Development and Research, Washington, DC.

Milburn, Norweeta, Earl Edwards, Dean Obermark, and Janey Rountree, 2021, Inequity in the Permanent Supportive Housing System in Los Angeles. Los Angeles: California Policy Lab. Accessed 16 January 2022 at <https://www.capolicylab.org/wp-content/uploads/2021/10/Inequity-in-the-PSH-System-in-Los-Angeles.pdf>.

Miler, Joanna Astrid, Hannah Carver, Wendy Masterton, Tessa Parkes, Michelle Maden, Lisa Jones, and Harry Sumnall, 2021, What treatment and services are effective for people who are homeless and use drugs? A systematic 'review of reviews,' PLoS One 16(7):0254729.

O'Flaherty, Brendan, 2019, Research on homelessness: A guide for economists (and friends), Journal of Housing Economics 44: 1-25.

---, 2022, Homelessness. In Klaus F. Zimmerman, ed., Handbook of Labor, Human Resources and Population Economics. Springer-Verlag.

Rolston H, Geyer J, Locke G, Abt Associates (2013). Final Report: Evaluation of the Homebase Community Prevention Program. New York: Department of Homeless Services.

Rosenheck, Robert, Wesley Kaprow, Linda Frisman, and Wen Liu-Mares, 2003, Cost-effectiveness of supported housing or homeless persons with mental illness, Archives of General Psychiatry 60 (September): 940-951.

Stergiopoulos, Vicky, Cilia Mejia-Lancheros, Rosane Nisenbaum, Ri Wang, James Lachaud, Patricia O'Campo, and Stephen W. Hwang, 2019, Long-term effects of rent supplements and mental health support services on housing and health outcomes of homeless adults with mental illness: extension study of the At Home/Chez Soi randomised controlled trial, *The Lancet: Psychiatry* 6: 915-925.

Tidderington, Emmy, 2019, Comment: Long-term effectiveness of housing and support services for homeless adults with mental illness, *The Lancet: Psychiatry* 6: 879-881.

Tidderington, Emmy, Amanda Ayakian, and Daniel Herman, 2022, Developing and implementation typology of Moving On initiatives, *Housing Policy Debate*, forthcoming

U.S. Department of Housing and Urban Development, 2021a, "HUD-Veterans Affairs Supportive Housing, HUD-VASH Vouchers, 2008-2020." Accessed 29 December 2021 at https://www.hud.gov/sites/dfiles/PIH/documents/HUD_VASH_Awards_%282008-2020%29_-_Public_Copy.pdf

---, 2021b, "PIT and HIC data since 2007." Accessed 29 December 2021 at <https://www.hudexchange.info/resource/3031/pit-and-hic-data-since-2007/>

Wong, Yin-Ling Irene, Trevor R. Hadley, Dennis P. Culhane, Steve R. Poulin, and Morris R. Davis, Brian A. Cirksey, and James L. Brown, 2006, Predicting staying in or leaving permanent supportive housing that serves homeless people with serious mental illness. U.S. Department of Housing and Urban Development, Office of Policy Development and Research, Washington, DC.

Table 1: Proportion of Days in Various Housing Statuses**Average days over length of the experiment**

Status	Treatment	Control	Difference
<i>At Home/Chez Soi</i>			
Emergency shelter	6%	16%	-10%
Street	3	8	-5
Subtotal, homeless (US)	9	24	-15
Temporarily housed	12	33	-21
Institutions	9	11	-2
Stably housed	73	32	+41
<i>HUD-VASH experiment</i>			
Homeless	14.5%	22.7%	-8.2%
Institutions	19.2	24.0	-4.8
Housed	66.0	52.9	+13.1

Notes: At Home/Chez Soi. Derived from text in Goering et al. (2014) on page 17 (days stably housed) and page 18 (days not stably housed). The sum of the categories for the treatment group is only 97%; the text does not explain why. This includes both treatment arms (Assertive Community Treatment and Intensive Case Management), and all five sites. It covers all days for 24 months of the intervention.

HUD-VASH experiment. Derived from table 2 (p. 945) in Rosenheck et al. (2003). "Treatment" is considered "Group 1: HUD-VASH" and "control" is "Group 3: Standard Care." The original table reports average days out of 90; I divided by 0.9 to get percentages. In original table, both groups in fact sum to 89.69 instead of 90. Participants were interviewed at 6 months, one year, 18 months, two years, and three years, and asked about housing in the previous 90 days. Numbers in table are averages over all these interviews. "Homeless" category in includes sleeping in "a substandard single-room occupancy hotel" (p. 942).

Table 2: Program Attrition Rates in Various Programs

Philadelphia PSH		HCV early		HCV late	
0-6 months	1.94%	0-10.8 months	2.66%	0-16.8 months	1.71%
6-12 months	2.63%				
12-18 months	2.35%	10.8-25.2 mon.	2.82%		
18-24 months	1.87%			16.8-51.6 mon.	1.17%
24-30 months	2.19%	25.2-68.4 mon.	1.60%	51.6-122.4 mo.	0.98%

Source: Wong et al. (2006) and McClure (2017). See text.

Figure 1: Capacity Expansion Scenarios

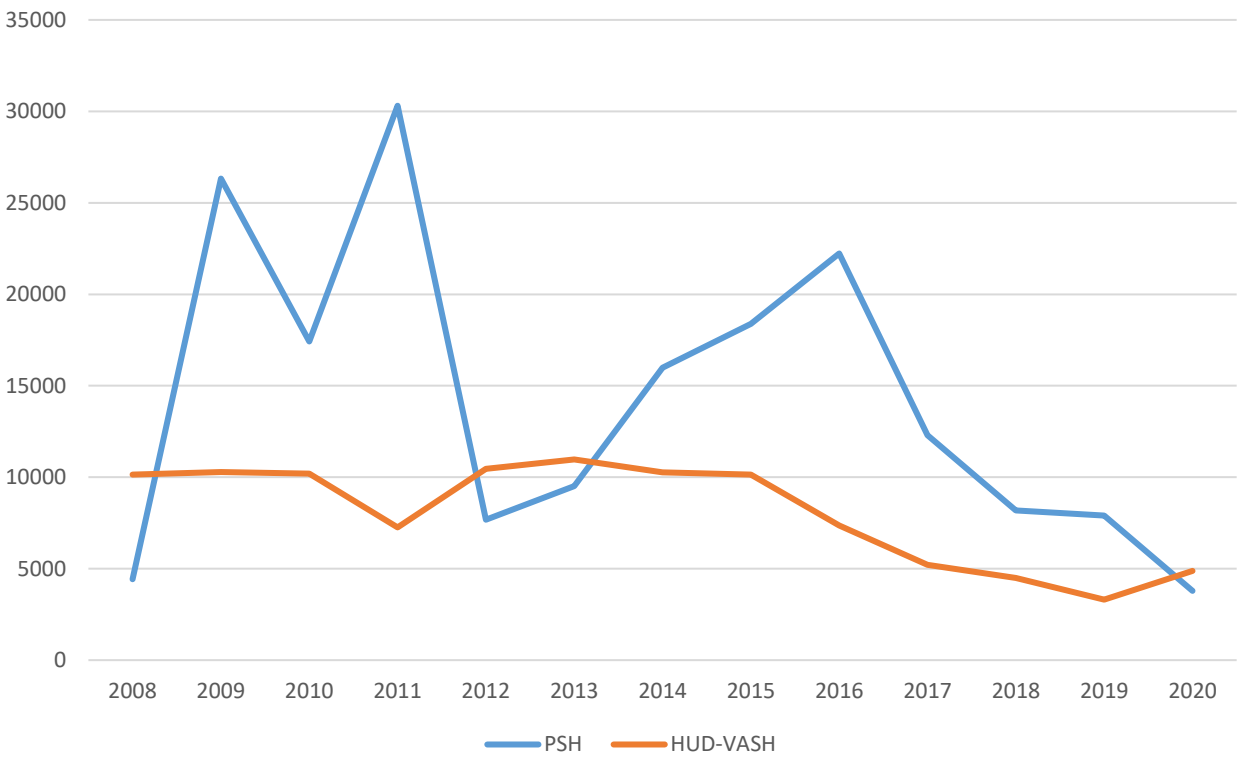


Figure 2: Reductions in PIT Count from At Home/Chez Soi ICM Simulations

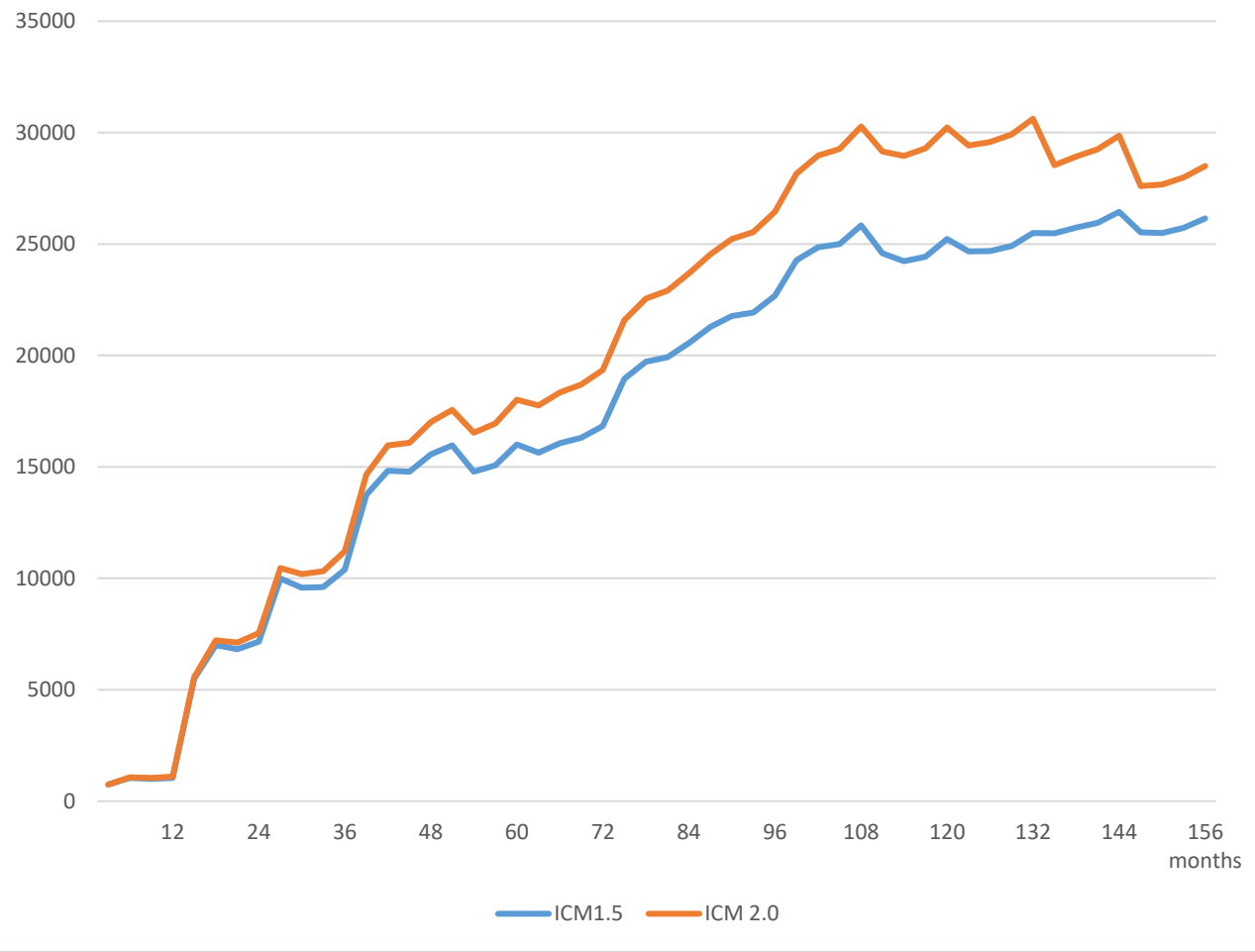


Figure 3: Reductions in PIT Count from At Home/Chez Soi ACT

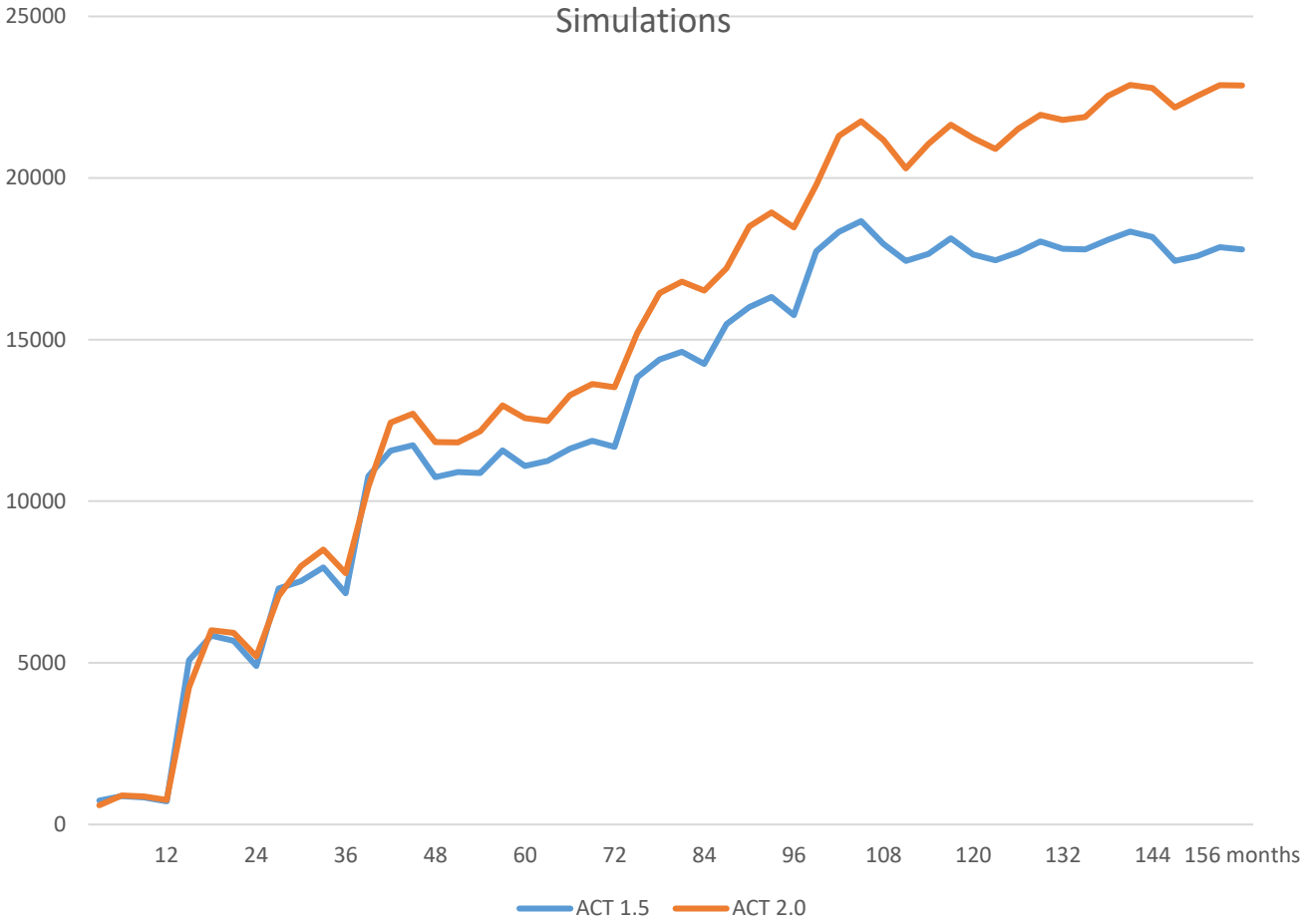


Figure 4: Reductions in PIT Count from HUD-VASH Simulations

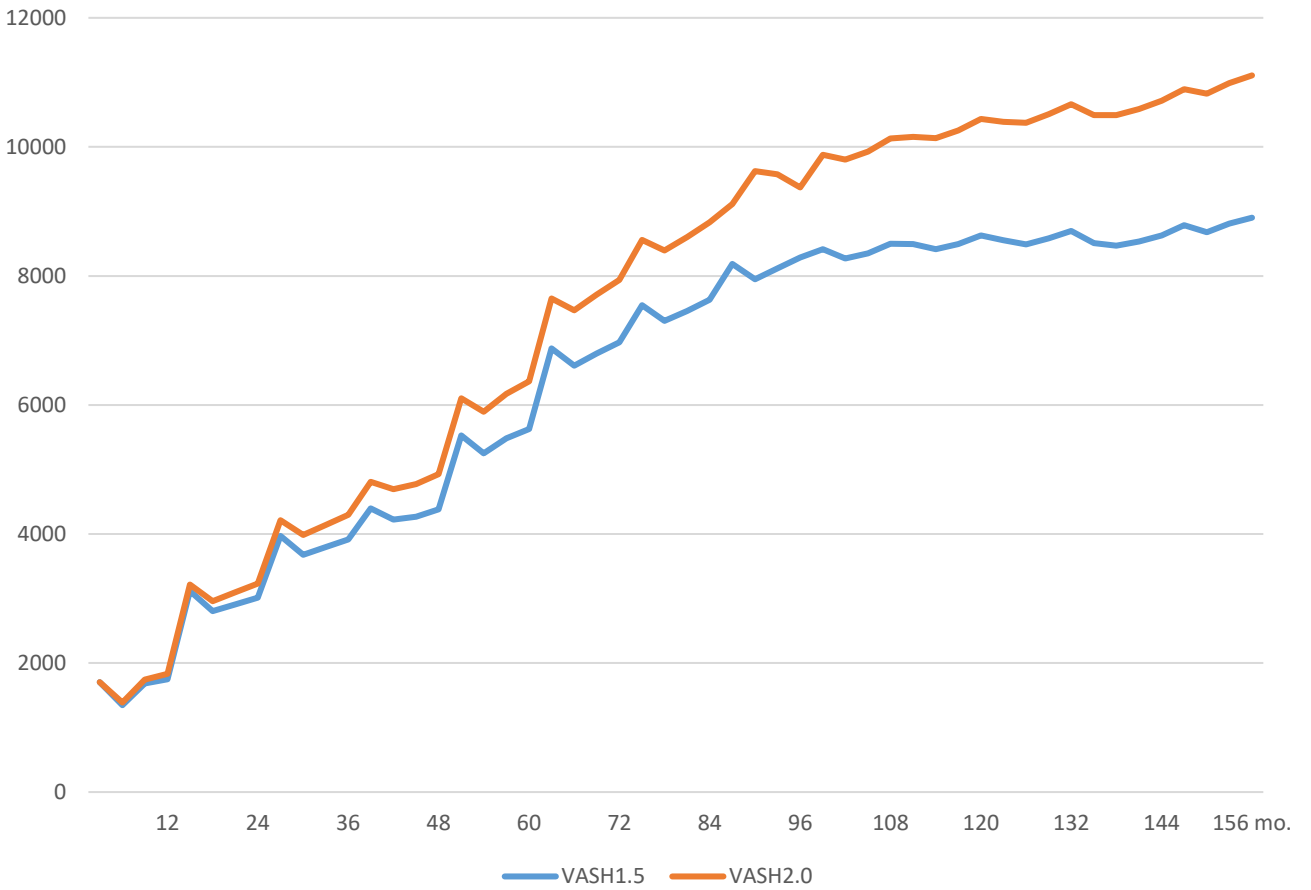


Figure 5: Ratio of PIT Count Decrease to Cumulative Capacity Expansion by Year

