Predicting and Preventing Homelessness in Los Angeles

TILL VON WACHTER, MARIANNE BERTRAND, HAROLD POLLACK, JANEY ROUNTREE and BRIAN BLACKWELL

UCHICAGO URBANLABS POVERTY LAB

CALIFORNIA POLICY LAB

SEPTEMBER 2019

CONTENTS

The California Policy Lab builds better lives through data-driven policy. We are a project of the University of California, with sites at the Berkeley and Los Angeles campuses.

The Poverty Lab works with government agencies and nonprofit organizations to identify effective tools for promoting social mobility and racial equity.

Summary

The California Policy Lab and the University of Chicago Poverty Lab have used County data on multi-system service use to predict homelessness among single adults receiving mainstream County services.[1] By identifying people at high risk of first-time homelessness or returns to homelessness and understanding risk factors associated with future homelessness, the County can more effectively target its homelessness prevention efforts to ensure limited resources are going to those most likely to benefit from them.

ACKNOWLEDGMENTS

This research was conducted in close collaboration with the Los Angeles County Office of the CIO, including Max Stevens, Fei Wu, Rachel Bryant, and Andy Perry; the Los Angeles Homelessness Initiative, including Phil Ansell, Elizabeth Ben-Ishai, Rowena Magana, and Bill Taylor. We thank them all for their support and guidance throughout. We are also grateful to a number of LA County officials who attended meetings, offered insightful feedback, or provided other crucial assistance, including Mark Greninger, Jose Soto, Maria Funk, Paul Arns, Sarah Mahin, Peter Lynn, Josh Hall, Meredith Berkson, Steve Rocha, Alex Devin, Mark Ghaly, Jonathan Liu, Cheri Todoroff, Elizabeth Boyce, Whitney Lawrence, Corrin Buchanan, Vanessa Martin, Sahelit Bahiru, Daniella Urbina, Taylor Schooley, Clemens Hong, Henry Bermudez, Kelvin Driscoll, Jenny Zogg, and Milli Tran, among others. We thank the broader research team at UCLA and University of Chicago including Janey Rountree (CPL), Brian Blackwell (CPL), Maya Buenaventura (CPL), Nefara Riesch (CPL), Nino Migineishvili (CPL), Carmelo Barbaro (UChicago Poverty Lab), Owen McCarthy (UChicago Poverty Lab), Emily Kristine Metz (UChicago Poverty Lab), Lingwei Cheng (UChicago Poverty Lab), and Halil Toros for their excellent work. Finally we thank Arnold Ventures, Scott Adelson, and the Max Factor Family Foundation in partnership with the Jewish Community Foundation of Los Angeles for their generous support. The views expressed are those of the authors and do not necessarily reflect those of the funders.

1 For the purposes of this project, "mainstream County services" include services provided by LA County departments reporting data to the Enterprise Linkage Project. Those departments include the Department of Health Services, Department of Mental Health, Probation, Sheriff's Department, Department of Public Health (Substance Abuse Treatment & Control), and Department of Public Social Services.

Background – policy context

On any given night, nearly 60,000 people experience homelessness in Los Angeles County,[2] and an estimated 141,000 are homeless in any given year.[3] In response to this growing crisis, voters in Los Angeles County passed Measure H, agreeing to increase their taxes to add an estimated $355 million in homeless services each year.[4] As reported in the 2018-19 Measure H 15-Month Report Card, 9,635 individuals entered permanent housing due to Measure H funding; 18,714 people entered crisis, bridge and interim housing funded in part or in whole by Measure H; 4,165 clients were linked to intensive case management services (ICMS); and about 3,300 have been assigned to either a federal or local rental subsidy for permanent supportive housing.[5] While the County has successfully navigated homeless individuals into available housing and other services, the homeless population continues to grow as inflow outpaces exits to permanent housing. In 2019, despite the influx of Measure H services, the homeless population in LA County (as measured by the Greater Los Angeles Homeless Count) grew by 12%.[6] Given the broader market forces driving housing costs and housing instability in Los Angeles, it is critical that the County and its research partners better understand the cause of inflows into homelessness and who is at highest risk in order to develop and test prevention strategies. It is also critical that the County rigorously evaluate services funded by Measure H so that County residents better understand what would have happened in the absence of these services.

Given the broader market forces driving housing costs and housing instability in Los Angeles, it is critical that the County and its research partners better understand the cause of inflows into homelessness and who is at highest risk in order to develop and test prevention strategies.

For the past two years, the California Policy Lab and the University of Chicago Poverty Lab ("the research team") have been working in close collaboration with the LA County Homeless Initiative and LA County Office of the Chief Information Officer (CIO) to predict homelessness among single adults receiving mainstream County services. The purpose of this work is to help identify people at high risk of homelessness and then leverage our growing understanding of risk factors to design and test homelessness prevention strategies. The research team has provided this analysis and research at no cost to the County.

2 2019 Greater Los Angeles Homeless Count, available at https://www.lahsa.org/documents?id=3423-2019-greater-los-angeles-homeless-count-los-angeles-County.pdf.

3 This figure is calculated using a combination of enrollment data in homeless projects from LAHSA's HMIS system, and the homeless flag in DPSS's data for General Relief (GR) recipients. Note that while individuals who are homeless in the HMIS are required to meet the HUD definition of homelessness, this is not a requirement to be flagged as homeless in the GR data.

4 "The Homeless Initiative," Los Angeles County, available at http://homeless.laCounty.gov/.

5 Measure H Citizens' Oversight Advisory Board Meeting Minutes, March 2, 2019, available at http://homeless.laCounty.gov/wp-content/uploads/2019/03/03.07.19-COAB-Mtg-Documents_FINAL2-2.pdf.

6 LAHSA: "Greater Los Angeles Homeless Count Shows 12% Rise in Homelessness." (June 4, 2019), available at https://www.lahsa.org/news?article=558-greater-los-angeles-homeless-count-shows-12-rise-in-homelessness.

In recognition of the importance of preventing homelessness, the Los Angeles County Board of Supervisors adopted a motion in May 2019 directing the LA County Homeless Initiative to collaborate with County departments, legal services, research organizations, and other experts to assess how to strengthen homeless prevention efforts within County departments. The Board directed the Homeless Initiative to include a description of the work being done by the research team on predicting homelessness and the implications of that work in the Interim Report. To that end, the research team is providing this summary of our work to date.

What we know about preventing homelessness

Experts note that homelessness prevention programs should be both **effective** and **efficient**.[7] **Effective** programs stop people at risk of homelessness from becoming homeless. **Efficient** programs target individuals and families who are at high risk of homelessness, i.e. those who would become homeless in the absence of assistance, rather than those who would find a way to maintain stable housing even without assistance. While there is very little existing research to help guide policy decisions on prevention, two recent studies in Chicago and New York offer reasons to be hopeful that prevention programs can be effective at preventing homelessness. However, effective targeting to ensure programs are efficient remains a challenge.

A prevention program managed by Catholic Charities in Chicago offered one-time cash assistance to families who called a hotline and self-identified as being at risk of homelessness. Callers demonstrating a minimum level of financial self-sufficiency and experiencing an eligible crisis qualified for one-time financial assistance up to $1,500. The program reduced shelter entry by 76% for program recipients when compared to a comparable control group who were eligible but happened to call on a day when funds were not available. While the program succeeded at reducing shelter entry, homelessness remained a rare outcome among this population: 99.5% of the individuals in the treatment group never entered shelter, compared to 98% of the control group. While this finding demonstrates that the vast majority of eligible callers were able to resolve their housing crisis by themselves, the prevention program was still cost effective because the cost savings to the shelter system exceeded the cost of running the program.[8] However, study authors noted that the program would be more efficient and cost beneficial if it were more effectively targeted to higher-risk callers.

7 Shinn, M. & Cohen, R. (Jan. 2019). "Homelessness Prevention: A Review of the Literature." Center for Evidence-Based Solutions to Homelessness. Available at http://www.evidenceonhomelessness.com/wp-content/uploads/2019/02/Homelessness_Prevention_Literature_Synthesis.pdf.

8 Evans, W. N., Sullivan, J. X., & Wallskog, M. (2016). The impact of homelessness prevention programs on homelessness. *Science*, 353(6300), 694-699.

A study in New York offers insight into how prevention services may be more effectively targeted to enhance their efficiency. The Homebase prevention program offers a variety of homelessness prevention service in community-based settings, including cash assistance, benefits counseling, case management, legal assistance, job placement, and other services. Shinn *et al*. (2013) developed and evaluated a screening model for families in New York City who applied to the Homebase program, though service providers could override the tool and exercise their own judgment. This model used demographic, employment, education, housing, disability, criminal justice history, domestic violence history data and other administrative data to predict risk of shelter entry for individuals who applied to Homebase.

An evaluation of Homebase found that during a 27-month follow-up period, Homebase reduced the average length of shelter stays by an estimated 22.6 nights when compared to a control group. The average number of nights in a shelter for all Homebase participants (including those with no nights in a shelter) was 9.6 nights and the average number of nights in a shelter for all individuals in the control group (including those with no nights in a shelter) was 32.2 nights. In addition, Homebase reduced the percentage of families who spent at least one night in a shelter from 14.5% to 8.0%.[9] Like the Chicago prevention program, the Homebase program was cost effective even though it had relatively modest effects. The evaluators of Homebase did, however, conclude that the program would have been even more effective had it been more efficiently targeted. Shinn *et al*. compared the families that the model identified as being at the greatest risk of homelessness with the families that Homebase program staff judged to be eligible for the program. As compared to program staff judgment, the Shinn *et al*. model had substantially higher precision (*i.e.*, correctly predicting shelter entry) at the same level of false alarms (*i.e.*, family that did not enter shelters in the absence of prevention services).[10] Greer *et al*. created a similar model to target individuals for Homebase. Greer *et al*. found that their model increased correct predictions by 77% (the model correctly predicted over 90% of shelter entry) and reduced missed cases of future homelessness by 85%.[11]

Both the Chicago and the New York programs demonstrate that short-term, relatively modest cash assistance and other temporary services can in fact prevent homelessness and reduce inflows by keeping individuals and families out of the emergency shelter system. That said, both programs also demonstrate the difficulty of efficiently targeting prevention programs. When a group of people all appear to be vulnerable, how do we know who is at highest risk of falling into homelessness?

9 Rolston, H., Geyer, J., Locke, G., Metraux, S., & Treglia, D. (2013). Evaluation of Homebase community prevention program. *Final Report, Abt Associates Inc, June, 6,* 2013.

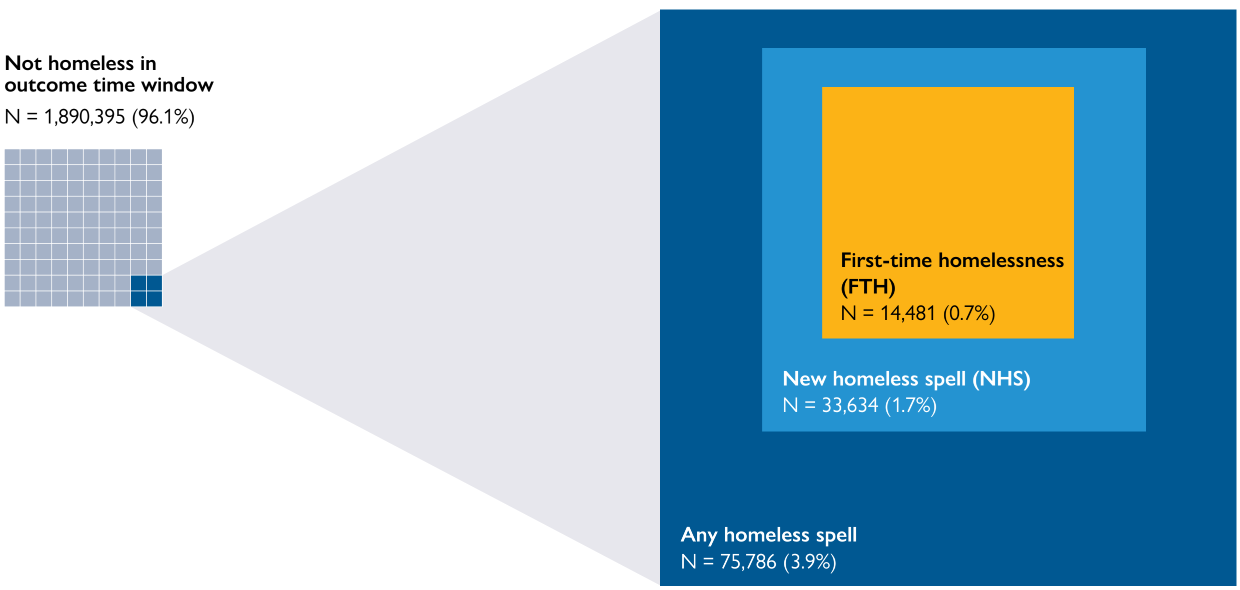
10 Shinn, M., Greer, A. L., Bainbridge, J., Kwon, J., & Zuiderveen, S. (2013). Efficient targeting of homelessness prevention services for families. *American journal of public health,* 103(S2), S324-S330.

11 Greer, A. L., Shinn, M., Kwon, J., & Zuiderveen, S. (2016). Targeting services to individuals most likely to enter shelter: Evaluating the efficiency of homelessness prevention. *Social Service Review*, 90(1), 130-155.

Research questions

In Los Angeles County, little is known about who among the millions of residents living in poverty are at high risk of homelessness. In any given year, the County interacts with approximately 1.9 million single adult clients through mainstream County services, including emergency room services; inpatient and outpatient medical treatment; mental health and substance abuse treatment programs; social safety programs like General Relief ("GR") and CalFRESH; and in the County jails and probation. In any given year, 76,000 single adults will experience homelessness, most of whom (~42,000) are experiencing an ongoing episode of homelessness continuing from the previous year. Roughly 34,000, however, are experiencing a new homeless spell, either returning to homelessness after being stably housed for at least six months (20,000) or experiencing homelessness for the first time (14,000) (see Figure 1).[12] The challenge facing homelessness prevention efforts is identifying who is at highest risk of experiencing a new homeless spell and connecting these individuals to services that effectively address their housing instability.

FIGURE 1. Homelessness among single adults in the ELP in calendar year 2017 (restricted to individuals with prior service histories)



12 These figures are restricted to single adults with prior County service history. When considering all single adults, 90,000 experience homelessness, with 48,000 of those individuals experiencing new homeless spells and 28,000 experiencing homelessness for the first time.

This research team's hypothesis is that advances in data science and **predictive analytics** can help identify who among those receiving mainstream County services is at greatest risk of homelessness. If we can efficiently target prevention resources, we can then use our understanding of risk factors to design more effective prevention strategies. These are the research questions that motivate our work:

- What are the key risk factors associated with future homelessness?
- Can predictive analytics help identify who is at risk of homelessness?
- What types of prevention programs reduce homelessness and for whom?

Methodology & data sources

Using Los Angeles County data,[13] the research team has developed a model for predicting homelessness in the County. The data sources for the project are derived from the Enterprise Linkage Project (ELP), which holds over 85 million service utilization records on 1.9 million single adults from seven agencies covering health services, benefits payments, law enforcement, and homeless services. The sheer scale of the data makes it ideal for the application of **predictive analytics**, which is the use of statistical models to make predictions about the future based on patterns and interrelationships between current and historical data. For example, for this research we predicted whether single adults experienced a new homeless spell in the 12-month outcome window of calendar year 2017, using data derived from calendar years 2012-16 as the prior service period.

Using predictive analytics, the research team has created models to predict two types of new homeless spells (NHS): **returns to homelessness** (RTH), in which the individual is not homeless in the six months prior to the outcome window, and **first-time homelessness** (FTH), in which the individual has no record of homelessness prior to the outcome window.[14]

13 The research team and LA County take data privacy extremely seriously and there are multiple measures in place to ensure that privacy. Individual County agencies participating in the Enterprise Linkage Project (ELP) run an encryption code that scrambles personally identifiable information such as names, birth dates, and social security numbers of the individuals in their data. The data is then uploaded to a secure server for inclusion into the ELP. The California Policy Lab has a data sharing agreement with the County CEO providing access to this de-identified data for the purposes of this project. The research team also used Homeless Management Information System (HMIS) data provided by the Los Angeles Homelessness Services Authority (LAHSA). The County encrypts the personally identifiable data in the HMIS using the same method that is applied to the rest of the ELP, and then shares the data with the research team. The research team does not have access to any information that would re-identify the individuals in the data set.

14 Because predictive analytics requires prior risk factors in order to make predictions about the future, only those County clients who have had interactions with County services prior to the outcome window (approximately 70% of individuals experiencing new homeless spells, and just over 50% of individuals experiencing first-time homelessness) can be included in the model.

Accurately measuring homelessness as an outcome is challenging. Individuals enrolled in homeless services in the Homelessness Management Information System (HMIS) must meet the definition of homelessness set by the Department of Housing and Urban Development (HUD), but relying solely on this measure risks underreporting the population experiencing homelessness. An alternative measure of homelessness is the homeless flag maintained by the Department of Public Social Services (DPSS) for GR clients. However, DPSS does not assess status using the HUD definition. In training the predictive models, we have used a combination of the HMIS and GR homelessness flags. Approximately 8% of all single adults who are flagged as homeless in these two systems are flagged in both, representing 24% of the HMIS-homeless and 11% of GR-homeless.

Results – can we predict homelessness in LA County?

Given its statistical rarity, it is difficult to predict homelessness. In fact, just 1.7% of approximately 1.9 million single adult County clients (33,600 people) experienced new homeless spells in calendar year 2017, of which 1.0% experienced a return to homelessness and 0.7% experienced first-time homelessness. To evaluate the performance of our model, we generated a list of people rank-ordered from highest to lowest risk of homelessness. The risk list can include any number of the highest risk individuals, depending on how it will be used. For example, the County could use a list of the 3,000 people at highest risk of first-time homelessness for a more focused, intensive intervention, or a list of 1% of people at greatest risk (19,600), which more closely approximates the actual size of the first-time homeless population in any given year. To offer options for various ways in which the County and its departments and agencies could use the risk list, we cut it into different sizes and assessed performance by calculating the **precision** of the lists, i.e. how many people on each size list actually became homeless?

For a risk list of the top 3,000 people at highest risk of experiencing any new homeless spells (including both first-time homelessness and returns to homelessness), 45.9% actually became homeless, including 27.1% in the HMIS. For a risk list of the top 3,000 people at highest risk of experiencing first-time homelessness in calendar year 2017, 33.5% became homeless, including 12.9% who were homeless in the HMIS according to the HUD definition.

Tables 1 and 2 report the precision of the models for lists of various sizes. They also report how much more likely the entire list, including people who didn't become homeless, is to experience homelessness compared to average County clients.

TABLE 1. **Model evaluation results for any new homeless spells among single adults, CY2017**

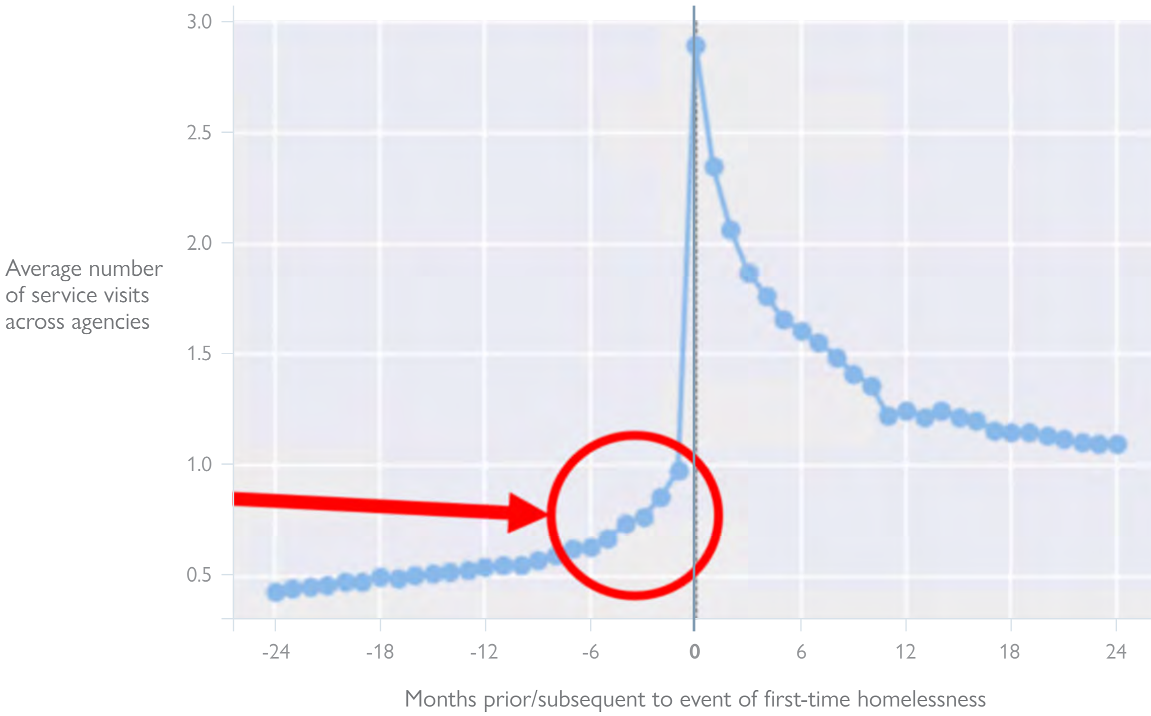
RISK LIST	PRECISION	RISK OF BECOMING HOMELESS COMPARED WITH AVERAGE COUNTY CLIENT
Top 3,000	45.9%	27 times more likely
Top 1% (N=19,600)	35.1%	21 times more likely
Top 2.5% (N=49,000)	28.6%	17 times more likely

TABLE 2. **Model evaluation results for first-time homelessness among single adults, CY2017**

RISK LIST	PRECISION	RISK OF BECOMING HOMELESS COMPARED WITH AVERAGE COUNTY CLIENT
Top 3,000	33.5%	48 times more likely
Top 1% (N=19,600)	23.6%	34 times more likely
Top 2.5% (N=49,000)	14.8%	21 times more likely

The phenomenon of homelessness is very time sensitive, with the immediate six months prior to the event of first-time homelessness containing, on average, a sharp spike in service usage (see Figure 2). Our analyses suggest that the predictive accuracy of the model in the field would be highly dependent upon the ability to continually refresh the model with data that is as current as possible.

FIGURE 2. Service utilization spike in six months prior to first-time homelessness (FY2013-17 data)



Who is on the homelessness risk lists?

It is worth noting that with these precision scores, some individuals on the risk list are "false positives", i.e. people who the model predicted are at risk but who did not actually experience a new homeless spell. However, all people on the risk lists, including the false positives, are vulnerable. The top 3,000 individuals who are at highest risk of experiencing first-time homelessness are **48 times more likely** to experience first time homelessness than average County clients. Those in the top 1% of the risk distribution are **34 times more likely** to experience homelessness in the next 12 months, and they are also intensive utilizers of services, with **nine times** as many arrests and jail bookings and **12 times** as many interactions with substance abuse prevention and control. People experiencing a new homeless spell are more likely to be male (69% vs. 54% of non-homeless County clients), and individuals experiencing homelessness for the first time are significantly younger than average (35 vs. 40).

What do we know about risk factors for homelessness?

In the prior five years of service history, 94% of those at risk of returning to homelessness[15] were receiving CalFresh; 86% were receiving General Relief; 88% had been in jail; 88% were Department of Mental Health (DMH) clients; 81% were Department of Health Services (DHS) clients; and more than 85% had contact with four or more agencies.

Much of our work thus far has been focused on maximizing the accuracy of the predictive models and performing descriptive analysis to understand who is at risk. The next phase of the project involves **interpreting** the predictive models in order to understand **risk factors for homelessness** – in other words, what are the underlying correlations in the data that have the most predictive power? What makes this next phase particularly challenging is the very large number of potentially important risk factors or 'features,' with almost 1,000 in the current model. Our current focus has been on the task of **feature selection** to determine which of these are most important.

We have found that a minimum of approximately 50 features is required to obtain acceptable performance in predictive models, with optimal performance requiring somewhere between 150 to 200 features. Although the most important features selected by the models tend to change frequently with the acquisition of new data and the application of new modeling techniques, some of the most important features include prior receipt of social safety net benefits, along with interactions with DHS, DMH, Sheriff, and Probation. The occurrence of multiple service types within one agency – for example, having both inpatient and outpatient visits – appears to be an important risk factor, along with temporal patterns such as the number of days since last visit (with shorter periods of time between visits suggesting higher risk). Certain geographical and spatial patterns are also emerging: having visited multiple service locations within one agency is a strong predictor, and certain zip codes appear to be at higher risk than others.

15 Because the risk lists that capture the top 1% of individuals at risk of a first-time homeless spell or new homeless spell is closest in size to the actual homeless population in any given year, the statistics below are all reported for the top 1% of the list.

Key insights

Our preliminary modeling results give us new insight into the individuals who are at risk of experiencing any new homeless spell, including first-time homelessness and returns to homelessness.

- The majority of single adults who will experience first-time homelessness or a return to homelessness are already clients of mainstream County agencies, which presents opportunities for intervention.
- Predictive analytics can greatly improve our ability to identify single adults at risk of homelessness and more precisely target prevention programs.
- Effectively serving the 1% of County clients who are at greatest risk of a new homeless spell would prevent nearly 6,900 homeless spells in one year.
- The 3,000 people at highest risk of first-time homelessness are 48 times more likely to experience first-time homelessness than average County clients.
- These clients are very vulnerable and are interacting with multiple systems, such as the mental health and criminal justice systems.
- Falling into homelessness happens very fast. The County and service providers must react quickly.

The insights gained from the modeling results can be used as guidance for existing and new prevention efforts. Stakeholders can:

- Proactively find people who are very high risk but who may not self-identify for services.
- Target scarce resources for people at highest risk.

Recommendations & next steps

The research team's goals in predicting first-time homelessness and returns to homelessness are (1) to efficiently target scarce prevention resources and (2) to test whether prevention programs are effective at reducing homelessness.

Improving efficiency: Over the coming months, the research team will continue to improve the precision of the models by adding data and working with LA County to optimize the pace at which the models are refreshed. The research team is also creating models to predict risk of homelessness among highly vulnerable populations, like older adults, single adults exiting the jail, or clients of mental health services. In these customized models, we can predict who among the subpopulation is at highest risk. The next step will be to work with County departments and other stakeholders to determine the best way to implement the risk models. Options could include (1) generating a high priority risk list for a multi-disciplinary problem-solving team, (2) generating customized risk lists for County departments, (3) creating customized risk lists by geography or population, and/or (4) testing the feasibility of creating a risk flag or risk score in County data systems.

Testing effectiveness: Given that tens of thousands of clients of mainstream County services are falling into homelessness each year, there is an enormous opportunity to leverage existing funding and service infrastructure to slow or halt a housing crisis before an individual becomes homeless. At the same time, there is very little research or evidence to help determine what level and type of assistance is needed to prevent homelessness. In its May 21, 2019 Motion, the Board of Supervisors highlighted the importance of assessing the efficacy of mainstream County systems in preventing homelessness. As the County mobilizes agency resources to prevent homelessness, we recommend that the County plan for evaluation of its prevention programs.

To that end, homelessness prevention experts have highlighted some high priority areas of research. Experts note that research on prevention programs for individuals leaving institutional settings (*e.g.*, correctional facilities or the foster care system) would be fruitful. Experts also note that research on housing subsidy programs could be particularly useful because this type of prevention has not been well-studied. In evaluating homelessness prevention programs, it is important to rigorously assess both effectiveness and efficiency and to not conflate the two. In other words, a homelessness prevention program that appears to be highly effective because enrollees do not experience homelessness in the outcome window might be inefficient if it targets people who are at very low risk. To differentiate between effectiveness and efficiency, evaluators need to measure outcomes against a counterfactual—what would have happened without access to the prevention program.[16]

The California Policy Lab and the University of Chicago Poverty Lab look forward to continuing to partner with County agencies to significantly advance this work in the coming months. If you have any further questions, please do not hesitate to reach out to Janey Rountree at janey@cpl.ucla.edu.

16 Shinn, M. & Cohen, R. (Jan. 2019). "Homelessness Prevention: A Review of the Literature." Center for Evidence-Based Solutions to Homelessness. Available at http://www.evidenceonhomelessness.com/wp-content/uploads/2019/02/Homelessness_Prevention_Literature_Synthesis.pdf.