

Homeless Veterans' Reintegration Program

Evaluation Design Pre-Specification Plan Item 11 – Addendum

The following final design decisions were made for the HVRP evaluation after the original design pre-specification plan was written:

- **Length of participant data panel and timeline.** Page 10 of the pre-specification plan states that we expect outcome data to be available for four quarters after program enrollment for the average participant. However, we were able to extend the timeline for the evaluation and collect additional follow-up data. We now expect to have eight quarters of data available after program enrollment for all study participants. This additional data collection also changed the planned timing for the release of the final report from fall 2022 to fall 2024.
- **Comparison group matching.** We are making two updates to the approach described on page 7 of the pre-specification plan for matching HVRP participants who received Wagner-Peyser services in addition to HVRP services to a comparison group of veterans experiencing homelessness who also received Wagner-Peyser services.

First, rather than matching HVRP participants to comparison group members who may have received Wagner-Peyser services in a different time period, we will now only match HVRP participants to comparison group members who received services starting in the same program year and quarter. Exact matching on program year and quarter of entry will ensure that all outcomes will be measured at the same time for both HVRP participants and comparison group members. This will be important given the large fluctuations in unemployment rates during our sample period caused by the COVID-19 pandemic.

Second, rather than matching HVRP participants to comparison group members directly within the same county of residence and who have the same program entry quarter, we will match HVRP participants to comparison group members within the same state who resided in counties served by at least one HVRP grantee. We will also match on county-level characteristics and include control variables for those characteristics in the impact analysis.

Matching directly on county of residence turned out to be impractical because many counties had only a small number of or no potential comparison group members. For example, for 13 percent of HVRP participants in our analysis, no potential comparison group members resided in the same county who received Wagner-Peyser services starting in the same quarter. For an additional 24 percent of HVRP participants, only one or two potential comparison group members in the same county began receiving services in the same quarter. Matching more broadly across the state will therefore allow us to include more HVRP participants in the analysis. This approach will also broaden the pool of comparison group members, likely yielding HVRP and matched comparison samples with better balance on demographic characteristics and pre-program employment and earnings.

At the same time, matching HVRP participants to comparison group members in different counties has the potential to introduce bias if comparison group members living in different counties have different non-HVRP services available to them or experience different job market opportunities. For example, HVRP grantees may be located in counties with additional employment counseling and job training programs available for veterans. To address differences in the availability of non-HVRP services we will match HVRP participants only to comparison group members in the same state who reside in counties served by at least one HVRP grantee.

To address differences in job market opportunities we will include the following county characteristics in the propensity score and impact estimation models (measured in the year before HVRP participants in our analysis began receiving services): poverty rate, unemployment rate, unemployment rate among veterans, labor force participation among veterans, number of veterans in the county, share of individuals living in urban areas, and the median household income for veteran families. We will assess balance on these county-level characteristics for the matched samples.

- **Propensity score estimation.** When constructing the comparison group for this analysis the goal is to identify a group of veterans experiencing homelessness who are likely to have similar labor market outcomes as HVRP participants, if the HVRP participants had not received any HVRP services. To achieve this goal, we will estimate propensity scores to measure how likely individuals are to receive HVRP services based on their demographic characteristics, pre-program employment and earnings, as well as the characteristics of their counties of residence. As described on pages 7 and 8 of the pre-specification plan, we will form a comparison group using those propensity scores and then assess balance on these baseline demographic, employment and earnings, and county characteristics to determine whether the comparison group is a good match for the HVRP group.

It may be that balance on the baseline characteristics could be improved if the propensity score estimation method included interactions between the characteristics. For example, including an interaction between gender and age could be more predictive of the likelihood of receiving HVRP services than including gender and age as variables by themselves. Because it would be difficult to manually check which of the many possible interaction terms provide additional predictive power, page 7 of the pre-specification describes two machine-learning algorithms that search for interaction terms: least absolute shrinkage and selection operator (LASSO) models and generalized boosted regression models (GBM). We conducted a proof-of-concept analysis of WIPS data from program year 2017 and found that both the LASSO and GBM approaches can produce good balance on observable characteristics between HVRP participants and the comparison group. However, the LASSO approach was quicker to implement computationally. We will therefore use the LASSO model as our primary method, and will use GBM if we are unable to achieve good balance after using LASSO.

The baseline characteristics and interaction terms that are predictive of whether veterans experiencing homelessness receive HVRP services may differ from the variables that are predictive of future labor market outcomes. The HVRP evaluation will therefore also use a double selection

LASSO to identify whether additional interaction terms are predictive of future employment outcomes (Belloni et al. 2014). If additional interaction terms that were not identified by LASSO as predictive of receipt of HVRP services are found to be predictive of future employment outcomes, we will add those interaction terms to the final propensity score model that we use for matching.

To determine whether we are achieving balance on baseline characteristics that are the most predictive of future labor market outcomes the study team will also examine prognostic scores when assessing balance. Using prognostic scores to assess balance can help reduce bias in studies using propensity scores (Stuart et al. 2013). Prognostic scores are constructed by using the control group to predict future labor market outcomes using the same covariates as in the propensity score models. The prognostic scores will be the predicted values of employment and earnings outcomes from this model, which we will calculate for both the HVRP and comparison groups. The study team will assess HVRP–comparison group balance on prognostic scores, along with the other important baseline variables such as pre-program employment and earnings.

- **Confirmatory earnings outcomes.** One of the primary goals of the HVRP evaluation is to determine whether HVRP has long-run impacts on participants' employment and earnings outcomes. The longest period for which we have follow-up data for the entire analysis sample is eight quarters, so page 8 of the original pre-specification plan specified earnings in the eighth quarter after program entry as the primary confirmatory earnings outcome.

However, after the pre-specification plan was written the COVID-19 pandemic caused substantial increases in volatility across quarters in labor market outcomes (Edwards et al. 2022; Smith et al. 2021). Higher volatility in quarterly earnings will lead to less precise measures of impacts on earnings outcomes. To mitigate the influence of increased volatility of quarterly earnings while still looking at long-run impacts of the HVRP program, we will use average earnings across the seventh and eighth quarters after program enrollment for our confirmatory (primary) earnings outcome. This will replace eighth quarter earnings as the confirmatory earnings outcome specified on page 8.

References

- Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. "High-Dimensional Methods and Inference on Structural and Treatment Effects." *Journal of Economic Perspectives*, vol. 28, no. 2, 2014, pp. 29–50.
- Edwards, Roxanna, Lawrence S. Essien, and Michael Daniel Levinstein, "U.S. labor market shows improvement in 2021, but the COVID-19 pandemic continues to weigh on the economy," *Monthly Labor Review*, U.S. Bureau of Labor Statistics, June 2022.
- Smith, Sean M., Roxanna Edwards, and Hao C. Duong, "Unemployment rises in 2020, as the country battles the COVID-19 pandemic." *Monthly Labor Review*, U.S. Bureau of Labor Statistics, June 2021.
- Stuart, Elizabeth A., Brian K. Lee, and Finbarr P. Leacy. "Prognostic Score–Based Balance Measures Can Be a Useful Diagnostic for Propensity Score Methods in Comparative Effectiveness Research." *Journal of Clinical Epidemiology*, vol. 66, no. 8, 2013, pp. S84–S90.