WAGE EFFECTS OF WIN NYC'S INCOME-BUILDING PROGRAM

By Ella Gottlieb, Milo Borek, Sofia Soares, Sophia Oh, Annika Shekdar, and Saragna Malladi

With contributions from Chris Mann, Ava Spiner, and Cassidy Teminsky

win

Abstract

Women In Need (Win) is New York City's largest provider of shelter for families with children experiencing homelessness, and part of their work involves programs that engender long-term income growth and foster financial independence among their clients. Within these efforts, WIN's Income Building Program (IBP) provides clients with educational, financial literacy, and employment upskilling services, and this paper seeks to quantify the impact of the program. The evaluation of the program was conducted by the Social Business Consulting (SBC) Club at Cornell University, and we applied multiple regression to establish the efficacy of the program depending on factors such as race, gender, age, family size, and education level. Additionally, we ran a return on investment analysis and discovered an ROI of roughly 2 using a variety of different methodologies, indicating net financial gain from investment and funding for the program. The findings imply a positive relationship between program involvement and income growth, and suggest that additional funding for WIN's Income Building Program will contribute to long-term economic gain.

Introduction

Resolving New York Citys homelessness crisis has been the mission of many shelter and supportive housing-focused nonprofits to help families experiencing homelessness, and Women in Need has sought the development of innovative programs that address the root causes of the issue. Win's Income Building Program is a multifaceted initiative offering services ranging from GED preparation courses to job fairs to financial literacy, and it helps support clients who are transitioning from unemployment to employment. Beyond clients who are unemployed, Win helps employed clients to increase their success through helping to increase their wages, transition from part-time to full-time work, obtain health benefits, and upgrade from off-the-books to on-the-books positions. In 2024, IBP had 3300+ program participants, 50+ young adults accepted to college, and 235 previously unemployed clients find employment through the program. While these metrics serve as a general gauge for impact, the data analysis conducted in this study sought to tangibly quantify the impact conferred to the participants and the other variables across client demographics and levels of program involvement.

In order to evaluate the tangible impact of this program, SBC looked into data collected by Win's IBP team in order to determine the causal effects generated by the program. Within this data analysis, we examined demographic information and client profile characteristics to evaluate the impact of IBP depending on a client's background. Additionally, the analysis delved into different metrics to quantify success, including outcomes like employment status, hourly wage, and annualized income. Finally, the data analysis concludes a positive return on investment and subsequent recommendation for investment into the program in order to further economic growth and reduce long-term utilization of social services.

Data Background & Preparation

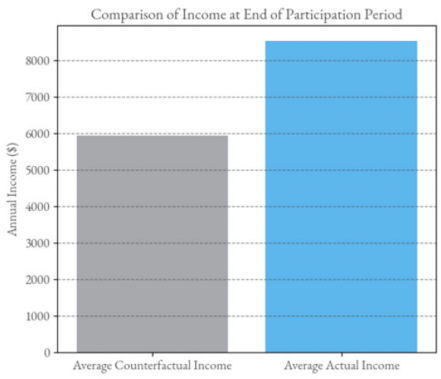
Win clients complete Income Building Assessments alongside their respective Income Building Specialist, with transitional housing clients completing the survey monthly and supportive housing clients completing it quarterly. In the survey, clients provide updates on their participation in the IBP as well as any changes in income and employment status. Three datasets were used: client demographics, which maps each client's unique ID to additional shelter, demographic, and medical history; along with both youth and adult responses to the Income Building Assessment. Variables include hours worked, salary, full vs. part time, benefits, and whether the client is working on or off the books.

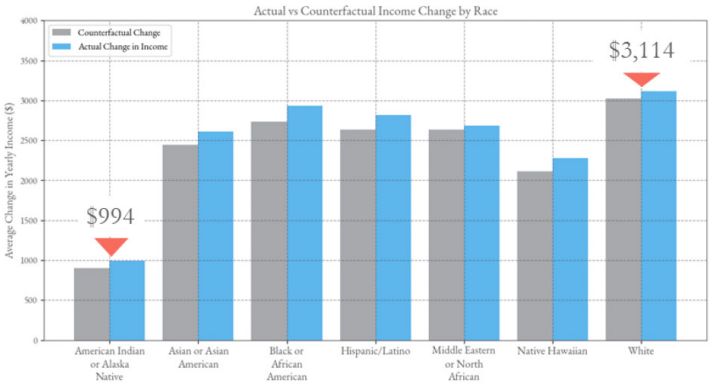
Due to the open-ended nature of the assessments, the data consisted of tens of thousands of unique survey responses, many of which included missing, duplicate, or outlier values. The Python libraries Pandas, Numpy, and Matplotlib were used for data cleaning, analysis, and visualization. We applied a four-step data cleaning process to the Income Building Survey datasets: we first removed all clients younger than 18 years of age and survey responses from before 2021 as that was Win's focus for the scope; we then dropped any columns that were empty or irrelevant to the economic analysis; next, we standardized several columns including Job Source and Race; finally, we removed rows in which a client was employed but did not record an hourly wage. We then removed "inactive" clients, or clients who had spent under 4 months in the program or completed the assessment under 4 times, since these cutoffs reflected minimal participation in the IBP. We did not use the youth responses to the Income Building Assessment because this caused the overall dataset to be extremely skewed.

Methodology Selection & Initial Analyses

To evaluate the impact of Win's IBP, we conducted a series of multiple regression analyses. While we initially considered alternative methodologies such as IMPLAN, Contribution Analysis, and Shift-Share Instrumental Variable techniques, we ultimately determined that multiple regression analysis would best leverage the client-provided data and yield the most accurate quantitative assessment of the IBP's impact. We also developed a bias guide documenting our assumptions about Win's survey data. Our bias guide addressed several forms of potential survey bias, particularly language bias and response bias, that could influence the interpretation of results.

Our preliminary exploratory analysis aimed to establish the program's overall impact on client's incomes. To achieve this, we compared clients' actual incomes by the end of their participation period to the counterfactual income they would have had if they didn't participate and just experienced the average NY 3% salary growth from their initial income. We found that clients who participated made on average $3,000 more than in the hypothetical situation that they did not participate, which was more than a 43% difference. Upon conducting an analysis of income change by race, we found that White, Black/African American, and Hispanic/Latino individuals saw the highest change in yearly income of more than $3,000 from the start to the end of their participation period, whereas American Indian/Alaska Native individuals saw the smallest benefits, with an average change of under $1,000.





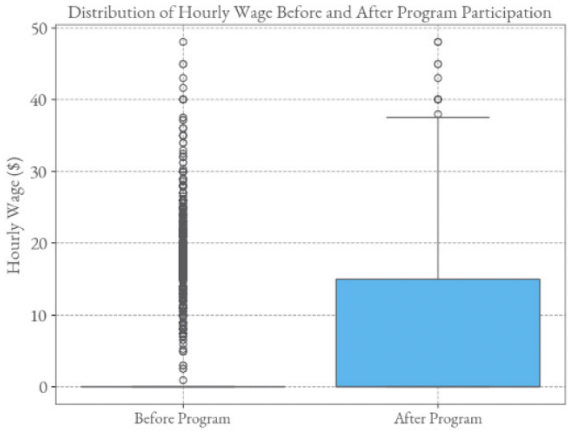
Linear Regression Introduction

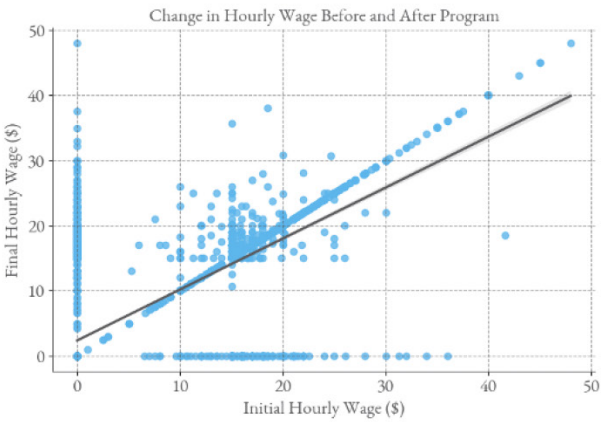
Linear regression is a statistical method used to model the relationship between a dependent variable and a single independent variable. It helps to predict the value of the dependent variable based on the value of the independent variable while assuming a linear relationship between the two. This method is widely used to assess the strength and direction of the relationship between the variables, providing insights into how changes in the independent variable affect the dependent variable. Linear regression is particularly useful in situations where there is a direct cause-and-effect relationship. Unlike multiple regression, which accounts for the influence of several predictors, linear regression focuses on understanding the effect of one predictor at a time, making it a simpler and more straightforward approach for analyzing relationships. Before conducting complex regression models, we ran several linear regressions to explore potential causation rather than mere correlation. Analyses included paired sample t-tests, multicollinearity tests, and other diagnostics to help us identify correlated variables and better isolate causal relationships.

Change In Hourly Wage Due To Program Participation

The first model evaluated the change in hourly wage due to program participation by comparing participants' initial and final hourly wages. The analysis showed an average wage increase of $1.53 after completing the program. The boxplot indicated that pre-program wages had clustering of low wages suggesting unemployed or very low-paying jobs and whiskers indicate wages before the program were highly dispersed. Post-program wage distribution showed a noticeable shift upward, with wages clustered in a higher range. Median wage is also significantly higher, distribution is less skewed, and wages are more uniformly distributed. This implies that IBP successfully helped participants move into higher-wage jobs and reduced variability in the data as concentration suggests predictable income range. The scatter plot confirmed a positive relationship between pre- and post-program wages. A majority of data points are above or along the regression line, indicating many participants experienced a wage increase, and the 45-degree line implies individuals see proportional or higher wage growth. Participants with low starting wages ($0–$15/hour) show a wider range of outcomes, suggesting varying levels of success depending on initial employment conditions. The scatterplot implies that IBP appears to be most effective in raising wages for low-income participants, and that there is overall positive wage growth, but higher-wage earners see less dramatic changes. To test whether these wage increases were statistically significant, both a paired t-test and a Wilcoxon signed-rank test were conducted. The results from both tests revealed extremely small p-values ($p \approx 0$), confirming that the increase in wages was highly statistically significant and unlikely due to random variation. Together, this evidence supports that IBP has a statistically positive impact on wages as participants saw wage increases that are unlikely to be due to random change.

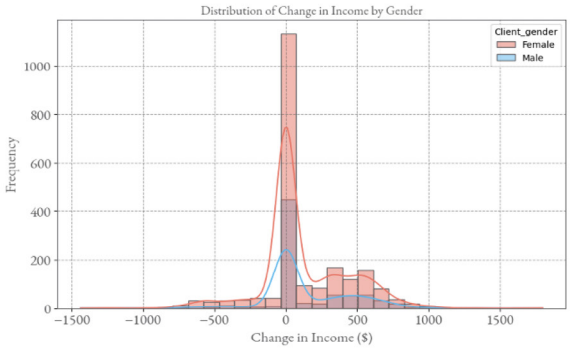
The analysis showed an average wage increase of $1.53 after completing the program.


Distribution of Hourly Wage Before and After Program Participation
Hourly Wage ($)
50
40
30
20
10
0
Before Program
After Program


Change in Hourly Wage Before and After Program
Final Hourly Wage ($)
50
40
30
20
10
0
0
10
20
30
40
50
Initial Hourly Wage ($)

Change In Income By Gender

This model analyzed the change in income by gender to assess whether male and female participants experienced different outcomes. This would help understand whether or not Win should allocate more of its funding towards one group over another or if there are significant demographic groups that can be addressed specifically. The box plot had a positive median, proving most participants saw an income increase after IBP. The median for males appears slightly higher than for females, and the distribution is fairly symmetric, implying income changes were relatively balanced. The histogram had a slight positive shift, indicating more people experienced an income increase than decrease. Male distribution has greater spread in both positive and negative wage changes, and males have a slightly higher proportion of large income increases. While the data showed that males had a slightly higher average income increase ($130) compared to females ($124), the difference was not statistically significant based on both a t-test and a Mann-Whitney U test ($p > 0.05$). These results suggest that the program benefited both genders similarly in terms of income growth, and no substantial gender disparities were observed in program effectiveness.



Length of Stay Compared to Change in Income

This model examined whether length of stay in the program was associated with change in income. This will help Win evaluate the longterm effects of IBP. The scatter plot shows a flat regression line, indicating no relationship between length of stay and income change. Staying in the program longer does not necessarily lead to income changes, and dispersed data indicates participants experience a range of changes, regardless of duration. A Pearson correlation coefficient ($r = 0.0079$) confirmed there was no significant linear association between the number of days participants stayed and their income growth. However, a Spearman correlation coefficient ($\rho = 0.1252$, $p < 0.05$) indicated a weak positive rank-based relationship, meaning that participants who stayed longer tended to have slightly better income outcomes, though the effect was minimal. Overall, length of stay alone does not meaningfully predict income gains for participants.

Return on Investment

Return on investment (ROI) is a calculation that gives a quantitative measure for the impact of an investment. In this case, the goal of an ROI calculation is to gauge the returns generated by the Income Building Program. For this calculation, we used the total program spend as the investment. We then calculated the return in several different ways. First, we used the income approach, which calculates the return as the change in annualized income from the first check in to the last check in of the year. This can be calculated by:

$$Income_{initial} = wage_{initial} * hours_{initial} * 52$$

$$Income_{final} = wage_{final} * hours_{final} * 52$$

$$\Delta Income = Income_{final} - Income_{initial}$$

$$\frac{\Delta Income}{Total\ Spend\ on\ program} = \text{ROI}$$

This method gave us a return on investment 2.08, which means that for each dollar invested in the income building program, 2.08 dollars are created in annualized income. It is important to note that this model does not imply causality, and some of the value created may be attributed to alternate means, such as common wage growth. However, the wage growth for IBP program participants is significantly higher than the wage growth for New York City as a whole.

For each dollar invested in the income building program, 2.08 dollars are created in annualized income.

The other method that we used to calculate the ROI was to normalize all salaries to 2021 levels. We did this with the intention of controlling for price growth, to calculate the change in purchasing power between the beginning and end of the program for participants. We did this using the Consumer Price Index (CPI), which is a normalized price of the typical urban consumer's market basket of goods. Doing this, we utilized the conversion formula to convert all wages and spending values into 2021 values:

$$Income_{2021} = Income \frac{CPI_{2021}}{CPI_{2024}}$$

Using this calculation, along with the method described above, we found a return of investment of 1.96. Again, we cannot attribute the change in purchasing power to only the Income Building Program, but this represents a significant change in purchasing power relative to the general population.

> *The wage growth for IBP program participants is significantly higher than the wage growth for New York City as a whole.*

Change in Employment

We wanted to determine the change in employment from when participants begin in the program to when participants end in the program. We only used data for program participants from 2021 to 2025. For any individual's first assessment during the selected time period, the probability of employment was 21.33%. For the final assessment during the selected time period, the probability of employment was 34.27%, indicating a change in nearly 13%. It should be noted that the numbers here are relatively low because many individuals included in the income building assessment are not the head of household and/or are not searching for employment.

Multiple Regression Introduction

Multiple regression is a statistical technique used to predict the value of a dependent variable based on two or more independent variable values. It allows researchers to examine the strength and direction of association among several factors and an outcome variable while controlling for potential confounding variables. By using several predictors in a single model, multiple regression distinguishes the independent contribution of each variable from the rest, providing a clearer and more nuanced analysis of what causes changes in the outcome variable. This method is particularly helpful in policy and decision-making fields, such as economics, education, and public health, which are influenced by numerous interacting variables. Unlike simple linear regression, which looks at the relationship between a single independent variable and one dependent variable, multiple regression looks at how various variables together impact a single outcome. In this paper, multiple regression is used to better understand the variety of demographic, education, and job-related variables influencing wage outcomes for participants to assist in program evaluation and future policy recommendations.

Multiple Regression Analysis

Hourly Wages

The research used an Ordinary Least Squares (OLS) model regression to determine the interaction between log-transformed hourly wages and demographic, educational, and employment attributes of working participants in Win's Income Building Program. The dependent measure was the natural log of hourly wage (Log_Hourly_Rate) because it standardizes wage data distribution. Independent predictors included age, work hours per week, work experience, gender, and education level (reference group: high school/GED). The goal of this model was to see how these variables predict wages in order to provide guidance for program evaluation and policy recommendations on employment readiness and education programs.

Education level was found to be the strongest and most reliable predictor of wages. Not having a high school diploma was associated with a 13.6% lower wage, and having a bachelor's degree was associated with a 6.8% higher wage than having a high school diploma or GED. Having more hours per week worked was also positively associated with wages, and age had a very slight but significant negative effect. Neither gender nor graduate degrees statistically significantly influenced wages when other variables were held constant. The model was significant in general, but it only accounted for a small amount of wage variance ($R^2 = 0.064$).

Even though the log transformation had normalized the distribution of the dependent variable, model diagnostics indicated that the residuals were not normal (as shown by the Omnibus and Jarque-Bera tests, high kurtosis, and skewness). This non-normality does not bias the coefficient estimates but can affect the validity of inference, including p-values and confidence intervals—especially in small samples or in making policy inferences based on statistical significance. Therefore, while the findings are directionally significant, caution must be exercised when interpreting the size of individual predictor effects. Future modeling may involve robust standard errors or other strategies to better address residual distribution issues.

OLS Regression Results

Dep. Variable:	Log_Hourly_Rate	R-squared:	0.064
Model:	OLS	Adj. R-squared:	0.064
Method:	Least Squares	F-statistic:	185.1
Date:	Tue, 22 Apr 2025	Prob (F-statistic):	5.21e-304
Time:	18:24:02	Log-Likelihood:	-442.78
No. Observations:	21623	AIC:	903.6
Df Residuals:	21614	BIC:	975.4
Df Model:	8		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	2.8172	0.011	252.854	0.000	2.795	2.839
Age	-0.0009	0.000	-5.052	0.000	-0.001	-0.001
Hours_Per_Week	0.0030	0.000	19.680	0.000	0.003	0.003
Work_Exp	1.593e-16	1.74e-17	9.166	0.000	1.25e-16	1.93e-16
Grouped_Bachelors	0.0678	0.012	5.858	0.000	0.045	0.091
Grouped_Graduate	-0.0199	0.020	-1.006	0.314	-0.059	0.019
No_HS_Diploma	-0.1361	0.008	-17.067	0.000	-0.152	-0.120
Education_Grouped_Other	-0.0580	0.008	-6.967	0.000	-0.074	-0.042
Some_College	-0.0463	0.009	-5.089	0.000	-0.064	-0.028
Client_gender_Male	-0.0045	0.004	-1.122	0.262	-0.012	0.003

Omnibus:	6240.310	Durbin-Watson:	1.921
Prob(Omnibus):	0.000	Jarque-Bera (JB):	70751.643
Skew:	-1.061	Prob(JB):	0.00
Kurtosis:	11.604	Cond. No.	9.19e+17

Yearly Wages

When calculating change in income, we wanted to analyze how income changed for those who started and ended with a job, to understand the influence of different potential Income Building Program actions for those who already have a job. For this regression analysis, we needed to only include those who seem to be done with the program (final check in was 6+ months since the last recorded assessment), and were consistent members of the income building program (at least 6 income building assessments). Using the dependent variable of log of the final annual income and a variety of independent variables, here are the results:

```
R-squared:                       0.982
Model:                             OLS
Method:                  Least Squares
F-statistic:                     172.7
Prob (F-statistic):          5.00e-272
Log-Likelihood:                 626.59
No. Observations:                  497

Significant variables (p < 0.05):

Log of initial wage:           t = 107.733, coef = 0.992
Online Job Profiles:           t = 4.242,   coef = 0.259
Email Address:                 t = 2.338 ,  coef = 0.075
Childcare:                     t = -2.400,  coef = -0.113
Business Link:                 t = -2.144,  coef = -0.401
Cover Letter:                  t = 3.026,   coef = 0.326
Needed to do Business Link:    t = -4.726,  coef = -0.419
Other Needs:                   t = 2.297,   coef = 0.070
```

Here, we can clearly see that the log of the initial wage is the most important variable in predicting the final wage for those in the income building program. This is to be expected, as nearly all participants will keep the same income level during the duration of their program participation. However, it is interesting to see that creating/managing online job profiles increased the wage of participants, who held jobs at a very statistically significant level. Additionally, cover letter assistance was also seen to be beneficial, with a statistically significant coefficient. On the other hand, needing childcare and a business link were both negatively statistically, meaning that those who had childcare assistance or a business link had lower final annualized income. This is likely due to alternate reasons outside the program, such as that those who needed childcare assistance were less likely to retain as high pay since they needed to take care of their kids.

This regression is extremely significant, as the R 2 value is .982, which means that this model explains 98.2% of the variation in the dependent variable.

Probit Model

We wanted to be able to predict the odds of employment for an income building program participant based on their first check in. To do this we needed to use a binary variable as our explanatory variable and therefore needed a model built towards that. The result of a probits model prediction is a z-score that can then be used to calculate a probability. In this model, we used the final employment outcome for those who are done with the program (those who have not had an assessment in 6 months) and who were consistent members of the program (those who had at least 6 assessments completed). We used final employment status as the dependent variable and a wide variety of dependent variables (all from the initial income building assessment to predict employment by the final check in.

```
No. Observations:                   2657
Model:                            Probit
Method:                              MLE
Pseudo R-squ.:                    0.6965
converged:                         True
Covariance Type:                nonrobust

Significant variables (p < 0.05):

Intercept:                          t = -8.025, coef = -2.599

Goals
Become employed:                    t = 2.440,  coef = 0.414
Off book to on books                t = 2.074,  coef = 0.936
Raise wages:                        t = 2.335,  coef = 0.602

Employed initially:                 t = 15.371, coef = 3.351
Number of check ins:                t = 3.956,  coef = 0.112
Job App assistance:                 t = 3.592,  coef = 0.619
Income Building Workshop:           t = -2.505, coef = -1.056
Needed to do Adult Literacy work:   t = -2.469, coef = -0.866
```

Here, we can see that the largest predictor of final employment is initial employment. However, those who had a goal from going from unemployed to employed had a large increase in employment percentage, as well as those who had employment goals of raising wages, and going from off books to on books for employment.

Additionally, those who had a larger amount of check-ins had a significantly higher percentage of ending up employed by the final check in, which shows the value of the individual check in. There were two variables that had negative effects on future employment, those were participating in an income building workshop and needing to work on literacy. Both of these are likely because of preexisting characteristics about individuals that made it so they needed to participate in the workshop and the literacy work. For example, someone who has lower literacy skills is less likely to get a job and more likely to need adult literacy work. In all, this result has an R 2 of .6965, which means that this model explains 69.65% of the total variation in the dependent variable.

Recent Policy Changes and Implications for Win

In recent years, a number of significant policy changes and initiatives related to assisted housing have been implemented that directly affect Win, specifically in the number of clients. As of March 2025, the New York City Department of Social Services (DSS) and the Department of Housing Preservation and Development (HPD) have revised the CityFHEPS (City Fighting Homelessness and Eviction Prevention Supplement) rental subsidy structure. Specifically, the Augmented CityFHEPS program—designed to subsidize rents above standard limits—will remain available through June 15, 2027. However, its use is now restricted to units priced at or below 120% of the 2024 Fair Market Rent (FMR). This policy change affects several key groups, including participants in the Housing Connect Lottery, individuals placed through Homeless Placement Service referrals, and new housing developments or units seeking to utilize Augmented CityFHEPS. While this adjustment is expected to yield cost savings for the City, it simultaneously constrains voucher holders' access to higher-quality housing, potentially limiting their housing choices. As a result, Win may experience increased demand for its services from individuals and families who are no longer able to secure higher-standard units previously within reach.

Additional policy shifts over the past few years have further transformed the assisted housing landscape. These include modifications to CityFHEPS eligibility criteria, employment requirements, and income thresholds for affordable housing programs. Notably, as of November 2022, CityFHEPS eligibility was broadened in several ways. The required work hours for qualifying families were reduced from 30 to 14 hours per week. Furthermore, tenants residing in single-room occupancy units are now required to contribute no more than $50 per month in rent, replacing the previous standard of contributing 30% of household income. Eligibility has also been expanded to include adults earning full-time minimum wage incomes at levels exceeding 200% of the federal poverty line. On a federal level, the U.S. Department of Housing and Urban Development (HUD) has made access to affordable housing programs more attainable by revising income eligibility thresholds. As of 2025, HUD now bases limits on 80% of the Area Median Income (AMI), rather than the former benchmark of 250% of the Federal Poverty Level (FPL).

Collectively, these developments increase the pool of individuals eligible for rental assistance and, by extension, for Win's support services. The reduction in required work hours and enhanced rental subsidies may enable clients to redirect time and financial resources toward other critical needs. This includes participation in Win's Income Building Program (IBP), which promotes financial literacy and employment readiness, thereby fostering long-term stability and self-sufficiency.

Policy Recommendations

In furthering the efficacy of WIN's IBP, government policy surrounding families experiencing homelessness in New York City and the nonprofits that look to service them is crucial. To better help nonprofits replicate innovative programs like WIN's IBP and to implement the most effective aspects of WIN into other shelter programs, there are policy changes New York City should implement. The city can work to improve the efficiency and scope of the CityFHEPS administration by increasing funding for the Housing Access Voucher Program, enabling more families to find permanent housing after shelter. To focus specifically on workforce development and long-term housing mobility, policy solutions include requiring employment counselors and income development programs in all family shelters as well as offering shelter-based educational resources. In addition, the city can dedicate funding streams for workforce development within transitional housing, which enables the shelter provider to use resources for job training, educational upskilling, and benefits counseling. Another policy initiative to consider is the subsidization or tax incentivization of employers who partner with shelter-based programs to seek employees, which would help programs like WIN's IBP to match clients with jobs. Collectively, the implementation of these policies would position New York City to create long-term alleviation to the homelessness crisis.

Data Constraints and Acknowledgements

Our analysis is limited by the availability of income-building assessment data, which is only collected from current Win clients residing in shelter. As a result, we lack information on individuals who have exited the program, likely leading to an underestimation of the full impact of Win's income-building services. This limitation is partially attributable to the benefits cliff effect, where even modest increases in income can result in the loss of public benefits. Once clients no longer require housing support, they typically disengage from income-building services and stop completing assessments. Consequently, our analysis does not capture long-term job retention or sustained employment outcomes following program exit.

During the data cleaning process, we removed records with substantial missing data for key independent variables and filtered out more "inactive" clients. In doing so, we assumed that this exclusion did not systematically bias the sample by disproportionately removing individuals from specific demographic groups.

Lastly, while the dataset did not meet all classical assumptions for multiple regression—specifically regarding multicollinearity and a leptokurtic distribution—we still report our results. The independent variables included in the model showed high levels of statistical significance, providing valuable insight despite these limitations.

Conclusion

This evaluation of Win's Income Building Program demonstrates that Win's shelter-based workforce programs like financial counseling, employment readiness, and education services result in meaningful improvements to participant employment and income growth, specifically in terms of income growth, job acquisition, and fiscal self-sufficiency. Through multiple regression, ROI analysis, and predictive employment modeling, our findings consistently show IBP's success—clients saw statistically significant pay increases, especially those with lower starting wages, and the program offered a strong return on investment, with each dollar invested generating nearly twice as much in annualized income. While some of these data limitations– client nonresponse following shelter exit and the effects of the benefits cliff–likely understate the program's full impact, conclusions strongly support sustained and expanded investment. Demographic and programmatic factors such as education level, digital job readiness, and sustained participation were also shown to influence outcomes, providing insights into how Win might continue to optimize their IBP. Cumulatively, these findings constitute a basis for continued support for IBP, as well as broader policy reform that institutionalizes workforce development services within New York City's family shelter system.

About Win

Since its founding to provide emergency shelter to four homeless women and their children, Win has grown into the largest provider of family shelter and supportive housing in New York City and the country. Win offers transitional housing and permanent supportive housing that's coupled with programs and services developed to support long-term housing stability. All of Win's services are guided by its mission—to transform the lives of New York City homeless families with children by providing the safe housing, critical services, and the ground-breaking programs they need to succeed on their own—so families can regain their independence and children can look forward to a brighter future.

Donate to Win

Follow Win's Advocacy Work

Follow Win

/winnyc.org

@winnyc_org

@winnyc_org

win www.winnyc.org