EXPLORING GENDER-SPECIFIC DIFFERENCES IN SUBSTANCE USE DISORDER RECOVERY CAPITAL: A MULTIPLE-GROUP LATENT GROWTH MODELING AND RANDOM FOREST APPROACH

by

Ginnie Sawyer-Morris

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Human Development and Family Sciences

Spring 2022

EXPLORING GENDER-SPECIFIC DIFFERENCES IN SUBSTANCE USE DISORDER RECOVERY CAPITAL: A MULTIPLE-GROUP LATENT GROWTH MODELING AND RANDOM FOREST APPROACH

by

Ginnie Sawyer-Morris

Approved: ____
Bahira Sherif Trask, Ph.D.
Chair of the Department of Human Development and Family Sciences

Approved: ____
Gary Henry, Ph.D.
Dean of the College of Education and Human Development

Approved: ____
Louis F. Rossi, Ph.D.
Vice Provost for Graduate and Professional Education and
Dean of the Graduate College

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____

Valerie A. Earnshaw, Ph.D.
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____

Allison Karpyn, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____

Heather Farmer, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____

Jennifer Carrano, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____

Federica Bianco, Ph.D.
Member of dissertation committee

ACKNOWLEDGMENTS

To my adviser and fellow committee members, Dr. Valerie Earnshaw, Dr. Allison Karpyn, Dr. Heather Farmer, Dr. Jennifer Carrano, and Dr. Federica Bianco, your collective mentorship over the past six years has positively impacted my development both as a professional and as a human. Thank you for your unending guidance and support.

To Jhane Campbell, Colleen Mueller, and the many women students in our lab, this research was made possible because of your hard work, dedication, and care. It was a privilege to work with you and I look forward to the positive change you will each affect in the lives of women and families throughout your lifetimes.

To my family and friends, but especially to Adam, thank you for holding me up on this journey. You reminded me that I was built for this work and held space for me when I needed it most. I am eternally grateful for the love you showed me over the past six years.

Finally, to the women and men who participated in this study, thank you for sharing your time and energy and giving a part of yourself to serve the broader recovery community. This work is dedicated to you.

TABLE OF CONTENTS

Appendix

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Recovery housing is a promising community-based treatment modality for the 21.2 million individuals living with substance use disorders (SUDs) in the United States. However, women and men face unique barriers in their recovery, and little is known about whether and how such barriers persist over time in recovery housing contexts. The current study sought to address this gap by identifying key, gender-specific predictors of women's and men's recovery status (i.e., stable versus unstable recovery) using a latent growth modeling and machine learning approach. Through secondary analysis of a community-based sample of Delaware sober living home residents, multiple-group latent growth modeling was used to capture gender-specific trajectories of women's and men's recovery capital. These trajectories were then used in a series of gender-specific random forest predictions to identify variables strongly associated with women's and men's recovery status. Findings suggest that while social support was the strongest predictor of both women's and men's recovery status, women presented with more trauma and co-occurring mental health disorders, made less money, and reported greater financial strain, stress, and depressive symptomatology compared to men. Given the gender-specific barriers women face in recovery, sober living homes represent an ideal context for the implementation of gender-responsive programming.

Chapter 1

INTRODUCTION

Gender may have important implications for women's and men's differential development of recovery capital while residing in sober living homes. Gender-specific differences are historically understudied in the substance use field, and while research in this area is growing, there remains a dearth of research on gender differences in community-based recovery contexts (e.g., sober living homes). Extant findings suggest that factors such as age, co-occurring mental health disorder diagnoses, and parenthood status differentiate women's and men's development of substance use disorder (SUD) as well as access to, and completion of treatment (Becker et al., 2017; McHugh et al., 2018). In order to establish an evidence base that delineates ways in which women's and men's recovery differs, it is important to look beyond the stage of abstinence initiation to later stages of recovery (e.g., early recovery) as well as to other contexts in which people recover.

Recovery housing is a promising mechanism of long-term recovery. Jason et al. (2020) estimated that 1.2% of individuals with SUD reside in recovery housing each year in the United States. Recovery housing refers to a range of housing models that create mutually supportive communities where individuals work towards improving their physical, mental, spiritual, and social well-being and gain skills and resources to sustain their recovery (National Council for Behavioral Health & National Alliance for Recovery Residences, 2018). Such resources are conceptually known in the SUD recovery literature as recovery capital (i.e., the resources a person has access to and can bring to bear to initiate and maintain recovery from substance use; Cloud & Granfield, 2008). While recovery residences vary widely in structure, all are centered on peer support and a connection to services that promote long-term recovery. Recovery housing benefits individuals in recovery by reinforcing a substance-free lifestyle and providing direct connections to other peers in recovery and recovery services and supports. While associated with improved social support (Polcin et al., 2017; Stevens et al., 2015), little is known about the ways in which recovery capital changes among individuals residing in sober living homes, and even less is known about how such trajectories differ by gender. The current study addresses this gap by examining the ways in which women's recovery experiences differ from men's using a gender-specific longitudinal analytic and machine learning approach. Findings are contextualized and interpreted in light of previous gender-specific research as well as relational-cultural theory. Results can inform gender-responsive SUD recovery policies and programming in the public health field.

1.1 Gender Differences in Substance Use Disorder, Treatment, & Recovery

1.1.1 Gender Differences in Substance Use Disorder

In the United States, more than 21.2 million Americans live with SUDs, which are chronic, relapsing health conditions characterized by substance use that is compulsive or difficult to control despite adverse consequences (National Institute on Drug Abuse, 2020; Hasin et al., 2013). While men have historically been two to three times more likely to develop an SUD, women's and men's SUD rates are beginning to converge (SAMHSA, Office of the Surgeon General [US], 2016). Despite this epidemiological shift, a large gap remains in our understanding of how women's SUD recovery differs from men's as well as the gender-specific factors that play a role. Available evidence suggests that motivational factors for substance use differ by gender, with men using substances primarily to activate the body's physiological reward system (i.e., for excitement, novelty) and women using substances in response to emotional cues, such as stress or depression (Greenfield & Grella, 2009; Haseltine, 2000; Liu & Kaplan, 1996; McHugh et al., 2018). Women are also more likely to initiate use later than men; however, once initiated, women are more likely to progress faster to dependence and then to treatment (McHugh et al., 2018), an effect referred to as telescoping.

1.1.2 Gender Differences in Treatment & Recovery

Substance use disorder treatment is intended to help individuals stop compulsive drug seeking and use (National Institute on Drug Abuse, 2018). Given that SUD is a chronic illness, treatment is often a long-term process that requires multiple interventions and regular monitoring. Treatment can take a number of forms (e.g., medication-assisted treatment, cognitive behavioral therapy), occur in a variety of settings (e.g., in-patient residential rehabilitation, recovery housing), and last for different lengths of time (e.g., 30 days, 90 days, > 1 year; National Institute on Drug Abuse, 2018). In contrast, SUD recovery is a dynamic, lifelong process of change through which people improve their health and wellness, live self-directed lives, and strive to reach their full potential (Laudet et al., 2002; National Institute on Drug Abuse, 2017; SAMHSA, 2012).

Women experience gender-specific barriers to their recovery at the treatment stage and beyond. In terms of treatment initiation, women are less likely to begin treatment due to factors related to lack of affordable childcare, limited services for pregnant women, fear of losing custody of children, more limited support for treatment, and fear of sexual harassment (Greenfield et al. 2007; Logan et al., 2006; Tuchman, 2010). Such barriers are likely to be more pronounced for women with histories of incarceration, women who experience homelessness, and/or women with histories of domestic violence (Bloom et al., 2003; Edwards et al., 2017). Women who do seek treatment are more likely than men to present with complex SUD diagnoses (e.g., co-occurring SUDs and mental health disorders [MHDs]; McHugh et al., 2018) and to have experienced traumatization during childhood (e.g., maltreatment, neglect) that often persists into adulthood (e.g., emotional abuse, intimate partner violence; Hecksher & Hesse, 2009; Lotzin et al., 2019). While in treatment, women experience more severe impairment associated with employment, social/family, medical and psychiatric functioning (Foster et al., 2016, Hernandez-Avila et al., 2004, McHugh et al., 2013, Wu et al., 2010).

Much less is known about gender differences beyond the treatment stage. Research examining recovery outcomes post-treatment in samples with both women and men suggest that social support is a key predictor of long-term recovery. Individuals in recovery derive social support from a variety of sources including peer support groups, recovery homes, as well as friends and family outside of their recovery community (McGaffin et al., 2018). Findings from sober living homes studies suggest that recovery homes may serve as a mediator to social support by strengthening access to a sober community with a shared mission (Stevens et al., 2015).

While very limited, gender-specific research examining recovery outcomes for individuals residing in sober living homes is growing. This line of research suggests that the differences that impact women's recovery during treatment persist into early recovery. Early recovery is a term often used to describe the stage in which a person is focused on achieving abstinence from substance use (Laudet et al., 2004); it is a sensitive period in which both treatment and relapse are likely to occur. Compared to men, women in sober living homes are more likely to report improved outcomes associated with relationships as well as persisting problems with psychological functioning (Grella et al., 2005). In women-only samples, findings suggest that women residing in sober living homes often experience mental health problems, financial strain, and housing insecurity (Edwards et al., 2017; Krentzman et al., 2022) and may derive particular benefit from a trauma-informed recovery environment (Edwards et al., 2021). Research also suggests that a sense of community matters for women in sober living homes and has been shown to predict posttraumatic growth and positive mental health symptoms (Edwards et al., 2021, 2022).

In response to a lack of supports for women in recovery residences, researchers have called for programming that takes the contexts and roles of women's lives into account. They have specifically cited a need for the integration of trauma-informed services (Cano et al., 2017; Edwards et al., 2017) as well as policies designed to address the financial and mental health needs of women in recovery (Krentzman et al., 2022). While this area of research is growing, it remains understudied. Available evidence points to a need for services that are responsive to the needs of women in recovery (i.e., gender-responsive). Although most sober living home organizations split houses by gender, they otherwise approach women's and men's treatment in similar ways. Thus, sober living homes currently adopt gender-specific but not gender-responsive approaches.

1.2 Evolution of Gender-Responsive Treatment Approaches

The terms *gender-specific* and *gender-responsive* are often used interchangeably in the literature to describe gender-based SUD treatment approaches (Grella, 2008). However, the terms have differential implications for SUD programming and policy. As of yet, there are no universally accepted definitions to distinguish between the two; however, it is generally recognized that *gender-specific* approaches refer to separate treatment programs for women and men (Prendergast et al., 2011). *Gender-responsive* programs can be mixed-gender or gender-specific but must be equally responsive to women's and men's treatment needs as well as to the social, economic, political, and cultural forces that shape the context of their lives. For women, this means a program should promote psychological growth, prosocial behaviors, and provide a secure environment to safely discuss histories of trauma, abuse, and addiction without fear of judgment (Bloom et al., 2003).

In order to adequately distinguish between gender-specific and gender-responsive treatment approaches, it is important to first acknowledge their historical context and evolving representations across the last several decades (see Figure 1.1).

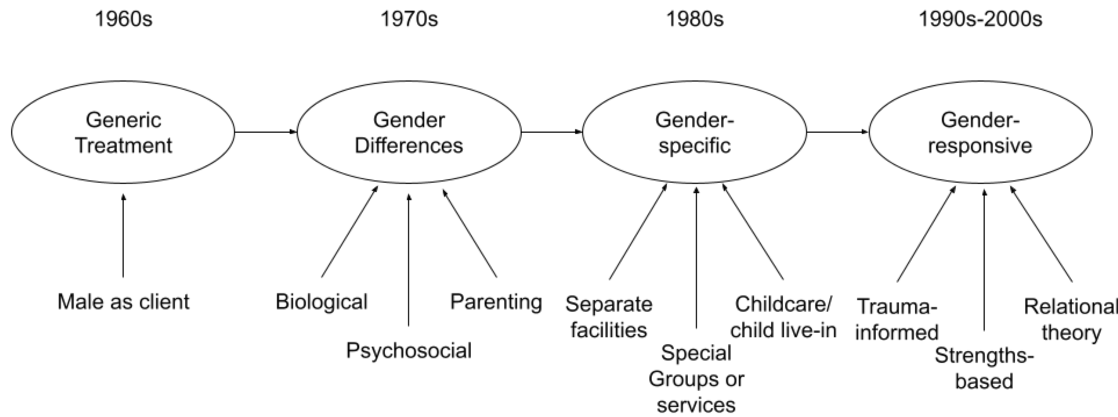

Figure 1.1 Evolution of Gender-based Treatment Approaches Adapted from Grella (2008)

In the 1960s, men were the primary clients for SUD treatment; women were viewed as supporters of men's sobriety, and not as individuals potentially in need of their own treatment (Grella, 2008). In the 1970s, the women's movement shed light on gender differences in social life; it was at this time that researchers and clinicians began to investigate sex/gender differences in women's and men's initiation and progression of substance use (e.g., differences in physiological effects) as well as differences in their treatment needs (Grella, 2008). Findings revealed that gender-specific barriers to treatment for women included lack of childcare, fear of stigma, and lack of family or financial support (Taylor, 2010). These challenges are not exclusive to women (i.e., men experience them as well), however, they are exacerbated for women due to overarching patriarchal structures that view women's health as synonymous with reproductive health. The investigation of these differences led to the development of separate, or gender-specific, treatment programs in the 1980s. Gender-specific treatment can be defined as treatment approaches that offer separate programs for women and men. Between the 1990s and 2000s, the concept of gender-responsive programming emerged (Grella, 2008). Gender-responsive programming “refers to programs where gender norms, roles and inequalities are considered, and measures are taken to actively address them. These approaches demonstrate a distinct focus on trauma-informed care as well as the central role relationships and intimate partners play (Covington, 2002; Grella, 2008).

In her decade review of changes in substance use treatment policies and programming, Grella (2008) discussed the evolution and various domains that characterize gender-responsive treatment programs. These include, but are not limited to: women-focused treatment services (e.g., prenatal/postnatal services, women-only groups in mixed-gender settings, parenting training/counseling, trauma/abuse counseling, women’s health services), children’s services (e.g., on-site childcare, live-in accommodations for children in residential settings, age and number rules regarding children’s participation, children’s mental health assessment, counseling/mental health services, children’s educational services, coordination with Child Welfare/Children’s Protective Services), and women-focused treatment orientation and processes (e.g., trauma-informed approaches, non-confrontational, empowerment, strengths-based, relational, developmental, trauma-informed, cultural competency).

1.3 Recovery Capital Theory

Over the past two decades, recovery capital (Granfield & Cloud, 1999) has emerged as a useful framework in operationalizing indicators of recovery. Conceptually, recovery capital represents the resources a person has access to and can bring to bear to initiate and maintain recovery from substance use (Cloud & Granfield, 2008; White & Cloud, 2008). In their theoretical framework, Cloud and Granfield

(2008) suggest that recovery capital can be grouped into four individual, yet overlapping, domains: human, physical, social, and cultural. The current study will explore human, physical, and social forms of recovery capital.

Human recovery capital includes a person's values, knowledge, educational/vocational skills and credentials, problem solving capacities, self-awareness, self-esteem, self-efficacy (e.g., self-confidence in managing high risk situations), hopefulness/optimism, perception of one's past/present/future, sense of meaning and purpose in life, and interpersonal skills. A person's *physical recovery capital* includes physical health, financial assets, health insurance, safe and recovery-conducive shelter, clothing, food, and access to transportation. *Social capital* encompasses intimate relationships and social relationships that are supportive of recovery efforts. Family/social recovery capital is indicated by the willingness of intimate partners and family members to participate in treatment, the presence of others in recovery within the family and social network, access to sober outlets for sobriety-based fellowship/leisure, and relational connections to conventional institutions (e.g., school, workplace, church, and other mainstream community organizations).

While recovery capital is a promising theoretical framework in the field of SUD recovery, key gaps exist. Of particular note is the issue of health inequity (i.e., inequitable access to resources that promote a healthy life because of social position or other socially determined circumstances; Centers for Disease Control and Prevention, 2020) and its potential impact on a person's ability to accumulate recovery capital. For example, the treatment-seeking challenges facing a middle-class married man who is employed full-time with benefits and has children differ from those of a single woman with children who lives below the federal poverty level and works full-time hours at multiple jobs but does not receive full-time benefits. The man may have more access to said treatment based on employment benefits, flexibility (related to caregiving and/or dual income), and/or insurance. By nature of being a lower-income single woman and mother, she is not distributed the same resources as her male counterpart and, consequently, has a potentially steeper hill to climb in her recovery (Bowleg, 2012; Rhodes & Johnson, 1997). These issues of inequity are not limited to gender but can extend to all aspects of a person's social position and identity (e.g., race, ethnicity, sexual orientation, ability). Inequitable access on the basis of social identity necessitates a critical analysis of the ways in which gendered societal and cultural power structures combine and operate as powerful mechanisms impacting women's and men's substance use and recovery. Recovery capital, on its own, does not take gender, or other forms of, inequity into account; therefore, when conducting gender-specific research, it is important to overlay this framework with gender-based theory (e.g., relational-cultural theory) that allows for the critical contextualization of women's and men's needs, roles, and contexts in recovery.

1.4 Recovery Housing

Recovery housing is an increasingly important community resource for individuals attempting to abstain from drugs and alcohol, especially for those in early stages of their recovery. Recovery housing refers to a "range of alcohol- and drug-free housing models that create mutually supportive communities where individuals improve their physical, mental, spiritual, and social well-being and gain skills and resources to sustain their recovery" (National Council for Behavioral Health & National Alliance for Recovery Residences, 2018, p. 2). Recovery residences are also referred to as sober houses, sober living homes/houses, sober living environments, recovery housing, and recovery homes throughout the literature and recovery field. Throughout this study, recovery housing will be referred to as sober living homes.

Increasingly, sober living homes have been recognized in the empirical literature for their ability to foster social support amongst residents. In their study examining recovery wellbeing, Cano et al. (2017) found that time spent in recovery homes was associated with increases in meaningful activities (e.g., employment, education, volunteering) and decreases in barriers to recovery. Given this and other evidence that stable housing in a sober community is positively related to the accumulation of recovery capital (e.g., economic and psychosocial resources and social support; Laudet & White, 2008; Polcin et al., 2016; Stevens et al., 2015), recovery homes represent a promising area for the implementation of gender-responsive policy and programming. In the state of Delaware, community-based organizations are working closely with state and local governments in an effort to establish guidelines for sober living homes as well as to create funding streams to support their continued operation. While efforts to create centralized standards for sober living homes are underway, very little is known about gender differences in recovery outcomes for women and men residing in sober living homes.

1.5 Current Study

The current study addressed this gap by exploring gender-specific differences in women's and men's development of recovery capital while residing in recovery homes. Cloud & Granfield's (2008) recovery capital theoretical framework is the foundation upon which the current study bases its design; however, interpretation of findings is guided by both recovery capital theory as well as relational-cultural theory (gender-based theory of women's psychology). The current study contributes to previous work on differences in women's and men's substance use recovery in two key ways. First, it moves beyond the stage of initiation in recovery to longitudinally examine the development of recovery capital during early recovery. Second, the study employs a gender-specific approach (i.e., analyzing women's and men's data separately) throughout every stage of the analysis and interpretation, allowing for findings that address the needs of both women and men in recovery.

1.5.1 Aim 1

The first aim is to examine differences in substance use recovery trajectories of women and men residing in recovery housing by estimating a series of gender-specific recovery capital latent growth models over time. This aim involves conducting a series of multiple-group latent growth curve models using longitudinal data from the Recovery in Sober Environments Study (RISE) and testing whether the intercepts and/or slopes of various recovery capital indicators differ significantly between women (n = 73) and men (n = 47) across time. Results may yield a more nuanced understanding of the differences between women's and men's recovery capital trajectories while residing in and transitioning out of recovery homes.

1.5.2 Aim 2

The second aim is to explore the relative importance of indicators of recovery capital and other conceptually relevant variables in predicting recovery status for women and men residing in and transitioning from sober living homes. The second aim utilized a machine learning approach (random forest). Because it is nonparametric, a random forest approach allows for a more integrated view of the importance of interdependent variables in predicting an outcome relative to traditional regression methods (e.g., latent growth curve), which assume linear correlation and are therefore limited by such dependencies. Random forests are capable of identifying, validating, and determining the predictive accuracy of factors from different domains (Hastie, Tibshirani, & Friedman, 2009), an important element when exploring indicators of recovery capital that are by nature distributed across a variety of domains (e.g., demographic characteristics, mental health, relationships, fulfillment of basic needs).

Chapter 2

METHODS

2.1 Procedures

Data were drawn from a longitudinal 10-month study that followed 120 individuals residing in Delaware-based recovery homes. The eight recovery homes included in this study are run by two community organizations. At the time of data collection, each home was staffed with a full-time, trained house manager and housed between seven and 16 residents at a time. During their stay, residents were required to remain substance-free, attend mutual aid meetings, and contribute to the upkeep of the residence. Researchers visited participating recovery homes monthly throughout the duration of the study until the COVID-19 pandemic began in March 2020, at which point data collection transitioned to fully remote.

Participants were recruited and enrolled on a rolling basis by research assistants between April 2019 - September 2019. Enrollment occurred during the researchers' monthly in-person visits to the recovery homes. All residents over the age of 18 who could read and understand English were eligible to participate. The study and its procedures were approved by the University of Delaware Institutional Review Board (IRB# 1256608-1). Participation was voluntary and informed consent was obtained from all participants. At the first visit, residents were asked to complete a baseline survey (~ 25 minutes) on tablets, or paper if preferred (< 5% of participants completed paper surveys). Nine follow-up surveys (~ 15 minutes each) were completed either in-person during the monthly visits or remotely via a survey link.

The remote option was offered to accommodate participants who were not able to attend the monthly visits due to work or other time conflicts as well as to provide an option for continued participation, should the resident move out of the recovery home during the study. Data collection lasted through August 2020.

2.2 Attrition

Both methodological and analytical approaches were used to address attrition. Methodological approaches are discussed here, and analytical approaches are discussed in section 2.4.1. Participant attrition is a common issue in longitudinal studies and an even greater challenge in populations with SUDs (Bootsmiller et al., 1998). Conversely, retention refers to a study's ability to maintain consistent engagement with participants throughout its duration. Recommended retention rates for substance use studies are lower compared to other healthcare populations, ranging from 60% to 80% (Digiusto et al., 2006; Hansten et al., 2000; Nemes et al., 2002; Polich et al., 1980; Scott, 2004). In a longitudinal study with multiple follow-up surveys, attrition introduces the risk of nonrandom systematic bias (i.e., likelihood that an underlying cause is related to the outcome predicting dropout). The concern with attrition in the current study was its potential relationship to recovery capital and/or recovery status; for example, a person who experiences a relapse or leaves the sober living home early (i.e., < 30 days) may be more likely to also drop out of the study. It is important to explore the extent to which attrition bias may compromise the results of a particular study, before concluding that attrition (on its own) invalidates the findings (Weisberg, 2005). As recommended by Mason (1999), extensive measures were taken to prevent attrition during data collection. These efforts included using incentives; employing a centralized tracking system, which was stored on a remote server (i.e., virtual private network) and updated monthly with participants' location and contact information as well as locator information for other contacts; including a remote option for completing follow-up surveys as an alternative to the brief in-person visits; and sending two text message reminders prior to each visit (or remote survey). COVID-19 introduced an additional and unforeseen risk for attrition bias. Steps taken to analytically explore the potential bias related to COVID-19 are detailed in section 2.4.1; Figure 2.1 provides a visual representation of the earliest time point in which a participant completed a survey during COVID-19.

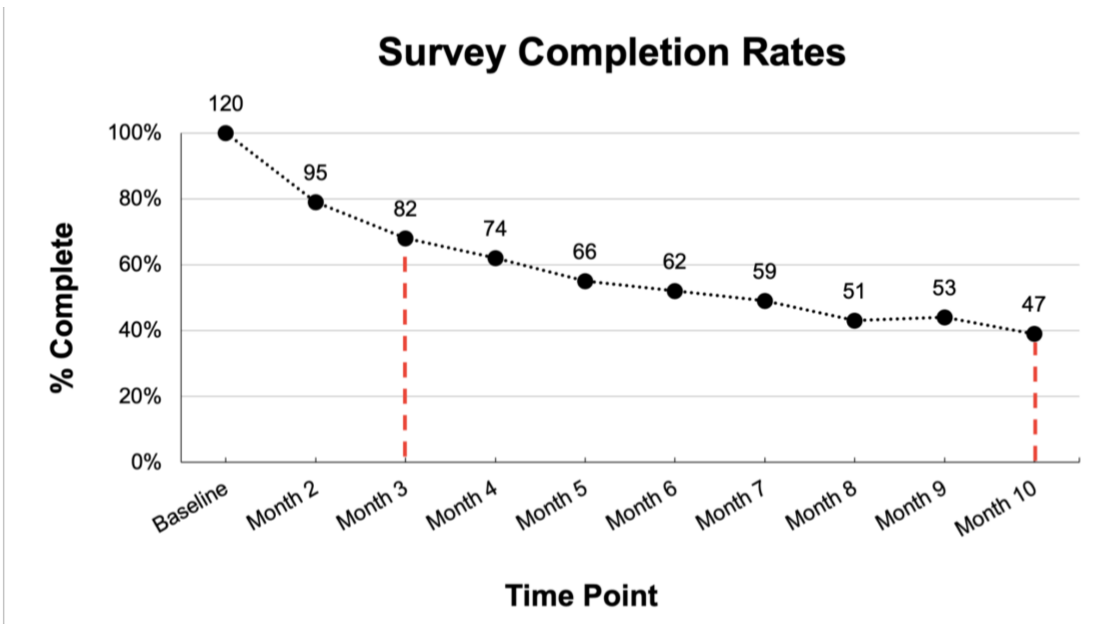

Figure 2.1 Survey Completion Rates across 10 Time Points.

Note. Survey completion rates are presented across 10 time points; data labels show the number of surveys completed at each time point. Red dotted lines at Months 3 and 10 represent the time points that were affected during COVID-19, with Month 3 representing the earliest time point where any of the participants completed a survey during COVID-19.

2.3 Measures

The following repeated measures were included in the study as indicators of recovery capital based on alignment with recovery capital theory: depressive symptomatology, mindfulness, perceived stress, internalized stigma, financial strain, and social support (see Table 2.1).

Table 2.1 Time-Varying Dependent Variables Nested within Recovery Capital Domains

Domains and Dependent Variables	# of items	Range of Scale	α	Example items	Scale Scoring
Human Capital					
Depression *(CES-D-10; Björgvinsson et al., 2013)*	10	0 (Rarely or none of the time) – 3 (Most or all of the time)	0.84 [b]	"I could not 'get going.' "	Sum
Mindfulness *(MAAS; Carlson & Brown, 2005)*	15	0 (Almost always) – 5 (Almost never)	0.94 [a]	"I find myself doing things without paying attention."	Mean
Perceived Stress *(PSS-4; Cohen et al., 1983)*	4	0 (Never) – 4 (Very often)	0.62 [c]	"How often have you felt that you were unable to control the important things in your life?"	Sum
Internalized Stigma *(SU-SMS; Smith et al., 2016)*	6	0 (Strongly disagree) – 4 (Strongly agree)	0.91 [a]	"I feel ashamed of having used alcohol and/or drugs."	Mean

Physical Capital					
Financial Strain *(Aldana & Liljenquist, 1998)*	8	0 (Never) – 4 (Always)	0.86 [b]	"I don't have enough money to pay my bills."	Sum
Social Capital					
Perceived social support *(MOS [+]; Sherbourne & Stewart, 1991)*	8	0 (None of the time) – 4 (All of the time)	0.96 [a]	"How often are the following types of support available to you if you need it… Someone to give you good advice about a crisis."	Mean

Note. Recovery capital domains (Cloud & Granfield, 2009) are presented in boldface. **Human Capital**: attributes that enable an individual to function effectively in society (e.g., knowledge, skills, education, health, mental health). **Physical Capital**: economic or financial capital (e.g., income, savings, property, investments, other financial assets). **Social Capital**: the sum of resources, actual or virtual, that accrue to an individual or a group by way of a social or relational network. CES-D 10 = 10-item Center for Epidemiological Studies Depression Scale, MAAS = Mindful Attention & Awareness Scale, PSS-4 = Perceived Stress Scale 4-item, SU-SMS = Substance Use Stigma Mechanism Scale, MOS = Medical Outcomes Study Social Support Survey. α = Cronbach's alpha reliability coefficient.

[a] α > .90 is considered excellent.

[b] α between .70 - .90 is considered good.

[c] α between .60 - .70 is considered marginal.

[+] Social support was measured using only the emotional/informational support subscale from the MOS survey.

2.3.1 Depressive Symptomatology

Depression symptoms were measured using the Center for Epidemiologic Studies Depression Scale Revised (CESD-R10; Björgvinsson et al., 2013; Radloff, 1977). Participants were provided a list of feelings and behaviors they might have experienced over the past week and were asked to indicate how often they felt this way using a four-point scale that ranged from "*Rarely or none of the time (less than 1 day)*" (0) to "*Most or all of the time (5-7 days)"* (3). Examples of items from the CESD-R10 include "*I was bothered by things that don't usually bother me"* and *"I could not 'get going'."* Raw scores were summed for each individual with higher scores indicating worse depression symptoms (any score equal to or above 10 is considered depressed). The summed scores at each wave were used as time-varying continuous indicators of depression in the multiple-group latent growth models (see section 2.4.2).

2.3.2 Mindfulness

The Mindful Attention & Awareness Scale (MAAS; Carlson & Brown, 2005) was used to measure the degree of mindfulness (i.e., awareness) to what is taking place in a given moment. The MAAS is a 15-item questionnaire with a six-point measure (0 = *Almost always,* 1 = *Very frequently*, 2 = *Somewhat frequently*, 3 = *Somewhat infrequently*, 4 = *Very infrequently*, 5 = *Almost never*). It is used to measure the frequency of mindfulness states, both attention to and awareness of, moment-to-moment experience and has been shown to possess strong internal consistency and construct validity across a variety of populations (Brown & Ryan, 2003; MacKillop & Anderson, 2007). Examples of items from the MAAS include "*I find it difficult to stay* *focused on what's happening in the present"* and *"I rush through activities without being really attentive to them."* Raw scores were averaged using the row mean (range = 0 - 4) with higher scores indicating greater levels of dispositional mindfulness. The average scores at each wave were used as time-varying continuous indicators of mindfulness in the multiple-group latent growth models.

2.3.3 Perceived Stress

Perceived stress was measured using the four-item Perceived Stress Scale (PSS-4; Cohen et al., 1983), a self-report measure of an individual's appraisal of general life stress. Responses to each item are rated on a five-point scale (0 = *Never,* 1 = *Almost never*, 2 = *Sometimes*, 3 = *Fairly often*, 4 = *Very often*). Example items include, *"How often have you felt that you were unable to control the important things in your life?"* and *"In the last month, how often have you felt difficulties were piling up so high that you could not overcome them."* The raw scores for these items were summed to obtain a total score (possible range 0-14), with higher scores indicating greater perceived stress. The summed scores at each wave were used as continuous indicators of perceived stress in the multiple-group latent growth models.

2.3.4 Internalized Stigma

The Substance Use Stigma Mechanisms Scale (SU-SMS; Smith et al., 2016) is an 18-item measure that assesses experiences of substance-related stigma. The items are divided into three subscales (six items per subscale): enacted stigma, anticipated stigma, and internalized stigma. Internalized stigma relates to how a person perceives themselves, rather than how others perceive them (e.g., anticipated stigma, enacted stigma); furthermore, it is possible that internalized stigma is a particularly pernicious barrier to recovery given that it plays a strong role in healthcare engagement (Earnshaw, 2020). Therefore, the internalized stigma subscale was used as an indicator of human capital. Example items from the internalized subscale include "*I feel ashamed of having used alcohol and/or drugs*" and *"Having used alcohol/drugs makes me feel like I'm a bad person."* Responses to each item were rated on a five-point scale (0 = *Strongly disagree,* 1 = *Disagree*, 2 = *Neither disagree or agree*, 3 = *Agree,* 4 = *Strongly agree*). Raw scores were averaged to obtain an overall mean score (range = 0 - 4), with higher scores indicating more stigma. The average scores at each wave were used as continuous indicators of internalized stigma in the multiple-group latent growth models.

2.3.5 Financial Strain

Financial strain was measured using the eight-item Financial Strain Survey (Aldana & Liljenquist, 1998). The items are divided into two subscales (four items per subscale): physical strain and meeting obligations. All items were included in the analysis. Responses to each item are rated on a five-point scale (0 = *Never,* 1 = *Rarely*, 2 = *Sometimes*, 3 = *Often,* 4 = *Always*). Example items include, *"Are you ever unable to sleep well because of financial worries?"* (physical strain) and *"I don't have enough money to pay my bills"* (meeting obligations). The raw scores for these items were summed to obtain a total score (possible range 0 - 32), with higher levels indicating greater levels of financial strain. The summed scores at each wave were used as continuous indicators of financial strain in the multiple-group latent growth models.

2.3.6 Social Support

The 19-item Medical Outcomes Study (MOS) Social Support Survey (Sherborne & Stewart, 1991) was used to measure social support. Participants were provided a list of items and asked to indicate how often support is available to them using a five-point rating scale (0 = *None of the time,* 1 = *A little of the time*, 2 = *Some of the time*, 3 = *Most of the time,* 4 = *All of the time*). The items are divided into four subscales: emotional/informational support, tangible support, affectionate support, and positive social interaction. Previous research suggests that emotional and informational support are associated with positive recovery outcomes such as treatment completion (Tracy et al., 2010), self-care (Brooks et al., 2017) and improved coping skills related to recovery (Tracy et al., 2010). Therefore, the emotional/informational support subscale (eight items) was included as an indicator of social capital. Example items from the emotional/informational support subscale include *"Someone you can count on to listen to you when you need to talk"* and *"Someone to give you good advice about a crisis."* Raw scores were averaged to obtain a mean score (range = 0 - 4), with higher scores indicating higher levels of social support. The mean scores at each wave were then used as continuous indicators of social support in the multiple-group latent growth models.

2.3.7 Recovery Status

Recovery status (i.e., unstable versus stable recovery) served as the dependent variable, or target, in the random forest models and was characterized by early exit from the recovery home (i.e., length of stay < 30 days in sober living home) and/or relapse. A number of steps were taken to construct the unstable recovery status variable (i.e., unstable). First, length of stay in sober living home (assessed at baseline) was calculated based on participant responses to the self-report item, "*For how long have you been living in the sober living home?*" Cases with < 30 days in-home at baseline were coded as 1 (unstable = 1; n = 63) and those with ≥ 30 days were coded as 0 (unstable = 0; n = 57). Next, of the cases with < 30 days in-home, those who moved out between the baseline assessment and month 2 were analyzed to determine whether their move-out date occurred within the < 30-day window. If so, these cases remained coded as 1 (unstable = 1; n = 15) and the remaining cases were recoded as 0 (unstable = 0; n = 105). Then, a relapse variable (participant experienced a relapse during study = 1; did not experience relapse during study = 0) was constructed based on responses to the time-varying self-report item, *"Have you used alcohol and/or drugs in the past 30 days or since your last interview?"* If a participant responded *"Yes"* to this question at any point throughout the study, they were coded as 1 (relapse = 1; n = 30). The relapse variable was then used to code the final step of the recovery status variable; that is, all individuals who experienced a relapse during the study were coded as 1 in the recovery status variable (i.e., unstable = 1, n = 37; unstable = 0, n = 83). Only three of the 37 participants with unstable recovery status both moved out of the home in less than 30 days and experienced a relapse at some point during the study.

The variable that resulted from these steps was a dummy indicator of unstable recovery where 1 = length of stay in home < 30 days and/or participant experienced a relapse and 0 = other. Less than 30 days was selected as a conservative indicator of early exit given the National Institute on Drug Abuse's (2020) recommendation for individuals to reside in recovery housing for at least 90 days to obtain maximum benefit (Polcin et al., 2010). Relapse was operationalized as a recurrence of substance use symptoms at any point during the study period. This approach to measuring relapse was appropriate because participants in the study were required to maintain sobriety while residing in the sober living homes. Relapse is defined as a recurrence of substance use symptoms after a period of abstinence or sobriety, often accompanied by resumption of pathological pursuit of reward or relief in response to dependence symptoms (Recovery Research Institute, 2022).

2.3.8 Demographic Characteristics

As part of the baseline survey, participants were asked to report their gender identity, age, race, ethnicity, LGBTQ status, education level, relationship status, employment status, parental status, incarceration history and parole status, primary substance of use, addiction severity, history of mental health disorder diagnoses, number of days in recovery, number of days in sober living home, number of relapses experienced over their lifetime, and sober living home affiliation (i.e., sober living home organization 1 versus organization 2). These variables were used to characterize the sample in the exploratory analysis and were included as input variables in the random forest models (see Table 2.2).

Table 2.2 Sociodemographic, Substance Use, and Mental Health Predictors

Predictors	# of categories/ Range	Categories (if applicable)/Scale	Variable Type	Description
Gender Identity	2	0 = Male 1 = Female	Binary Categorical	Time-invariant indicator of Gender Identity. *
Age	19-64	N/A	Continuous	Self-reported Age at Baseline.
Race/Ethnicity	5	1 = White, non-Hispanic 2 = Black, non-Hispanic 3 = Hispanic 4 = Native American, non-Hispanic 5 = Other	Categorical	Time-invariant indicator of Race/ethnicity.
LGBTQ	2	0 = Not LGBTQ 1 = LGBTQ	Binary Categorical	Individuals who identified as Transgender, Gay or Lesbian, Bisexual, or Other were coded as LGBTQ.

Education level	1	1 = Some HS 2 = HS Diploma/GED 3 = Some College 4 = 2-yr. College Degree 5 = 4-yr. College Degree 6 = Some Grad School 7 = Grad School Degree	Ordinal	Ordinal indicator of highest level of education at baseline.
Relationship Status	7	1 = Single 2 = Dating 3 = Married 4 = In a relationship 5 = Separated 6 = Divorced 7 = Widowed	Categorical	Categorical indicator of relationship status at baseline.
Employment Status	2	0 = Not Employed 1 = Employed Full-time/Part-time	Binary Categorical	Dummy indicator of employment status at baseline.
Parental Status	2	0 = No children 1 = Children	Binary Categorical	Dummy indicator of parental status at baseline.
Incarceration History	2	0 = Never been in prison/jail 1 = Experience in prison/jail	Binary Categorical	Dummy indicator of incarceration history at baseline.

Current Parole Status	2	0 = Not on Parole 1 = Currently on Parole	Binary Categorical	Dummy indicator of parole status at baseline.
Primary Substance of Use	4	1 = Alcohol 2 = Cocaine 3 = Opioids 4 = Other	Categorical	Categorical indicator of primary substance of use at baseline.
Addiction Severity *(LDQ; Raistrick et al., 1994)*	3	1 = mild (0-20) 2 = moderate (21-25) 3 = severe (26-30)	Categorical	Sum composite of LDQ was calculated; scores were used to create categorical variable reflecting the normative addiction severity ranges for the LDQ.
Number of Mental Health Disorder Diagnoses	7	0 = 0 Diagnoses 1 = 1 Diagnosis 2 = 2 Diagnoses 3 = 3 Diagnoses 4 = 4 Diagnoses 5 = 5 Diagnoses 6 = 6 Diagnoses 7 = 7+ Diagnoses	Ordinal	Ordinal indicator of number of lifetime mental health diagnoses reported at baseline.
Time in recovery (days)	1	4 – 5,475	Continuous	Number of days in recovery reported at baseline.

Number of Days in Sober Living Home	1	1 - 1,260	Continuous	Number of days in sober living home reported at baseline.
Number of lifetime relapses	3	1 = 0 Relapses 2 = 1-2 Relapses 3 = 3+ Relapses	Categorical	Categorical indicator of number of relapses (i.e., times participant started using drugs or alcohol after a period of sobriety) reported at baseline.
Sober Living Home Affiliation	2	0 = Organization 2 1 = Organization 1	Binary Categorical	Dummy indicator of sober living home organization affiliation where 1 = Organization 1.

Note. LGBTQ = Lesbian/Gay/Bisexual/Transgender/Queer or Questioning. α = Cronbach's alpha reliability coefficient.
* Participants who identified as Transgender Male-to-Female were coded as Female; participants who identified as Transgender Female-to-Male were coded as Male. LDQ = Leed's Dependence Questionnaire.

2.4 Analytic Approach

Gender-specific methods were used to address the primary aims of the study, which were 1) to determine whether women and men differ significantly in initial levels (i.e., intercepts) and rates of change (i.e., slopes) of recovery capital over a six month period using a traditional multiple-group latent growth modeling analysis, and 2) to explore the relative importance of recovery capital and other key variables (e.g., sociodemographic, substance use, and mental health) that influence women's and men's recovery status using a machine learning approach.

Before addressing the primary aims of the study, a comprehensive attrition analysis was conducted to explore patterns of missingness and potential bias. Next, descriptive, bivariate, and exploratory data analyses (e.g., outlier analysis, checking distributional assumptions, chi-squares, independent samples t-tests) were performed using Stata (version 15.1). To address the first aim of the study, unconditional multiple-group latent growth models were estimated to elucidate the gender-specific differences that potentially exist between women's and men's incremental and decremental development of recovery capital and its impact on recovery outcomes. Models were estimated using the R package lavaan (Latent Variable Analysis - version 0.6-9; Rosseel, 2012). Finally, information (i.e., parameter estimates, residuals, goodness of fit statistics) gained from the multiple-group latent growth models in the first aim were used as features in three random forest models, which together addressed the second aim (i.e., what variables influence women's and men's recovery status).

2.4.1 Attrition Analysis

Once data were collected, potential bias was addressed by implementing a comprehensive, multi-stage missing data analysis where patterns of missingness were 1) identified quantitatively, 2) coded using domain- and sample-specific knowledge, 3) tested for internal validity against information documented in the data collection tracking sheet, and 4) re-tested for mutual exclusivity. The "misstable" package in Stata 15.1 was used to quantitatively identify patterns of missingness in the time-varying recovery capital outcome variables (six time points). This method of exploration provides data-driven information about missing data (i.e., number of patterns; see Appendix A for an example of a missing data table using the time-varying social support variable). Next, domain and sample-specific knowledge were used to explain the patterns and develop categories, which included: "100% completion," "lost to follow up," "deceased," "electively dropped from the study," "item-level missing," "one wave of data," and "sporadic completion" (see Table 2.3 for a description of each category). Dummy codes were created to quantitatively characterize each of the codes. Mutual exclusivity (i.e., categories account for all patterns of missingness across all participants) was tested using the "fromdummies" package in Stata 15.1. This package uses a series of user-specified dummy variables to create a categorical variable with mutually exclusive categories. If the dummy variables do not represent mutually exclusive categories, the package will return an error. Once mutual exclusivity was validated, the dummy variables were used as binary categorical features in the random forest classification models as a way to control for the effect of missingness.

Table 2.3 Final Missing Data Categories

Missing Data Category	*n*	Description
Deceased	2	Participant died during the course of study as a result of overdose.
Electively Dropped	2	Participant electively dropped out of the study.
One Wave of Data*	17	Participant completed only one wave of data.
Item-level Missing	17	Participant completed all surveys but skipped items at one or more waves.
Lost to Follow Up	25	Participant was lost to follow up.
Sporadic Missing Across Time	26	Participant missed one or more follow-up surveys but resumed completing at a subsequent wave.
Observed (100% Complete)	31	Participant completed all data across all variables and all six waves.

Note. * $n = 6$ cases satisfied conditions for both "Electively Dropped" and "One Wave of Data;" however, it was determined that completing one wave of data poses a greater threat to validity compared to electively dropping. Therefore, individuals who electively dropped after completing only one survey were grouped in the "One Wave of Data" category.

Once patterns of missingness were identified, attrition bias was explored using a logit analysis. Loveland and Driscoll (2014) found that, in the U.S., 75% to 80% of SUD treatment-seekers disengage at one of the multiple stages of the enrollment and treatment process. Because the response rate (i.e., retention) at month six (52%; see Figure 2.1) was eight percentage points below the minimum recommended retention rate for substance use studies (60%; Hansten et al., 2000), a more conservative measure of completion was used to test bias. A dummy variable, "conservative," was created where 1 = individuals with 100% completion across all variables and time, and 0 = any pattern of missingness. Key independent variables at baseline were used to predict likelihood of membership in either category using a logit analysis. Because COVID-19 introduced an additional and unforeseen risk for attrition bias, a variable indicating whether participants completed surveys during COVID-19 was included as a predictor in the logit analysis. Results from this analysis provided information about the extent to which individuals with missing data differ from individuals with complete data; findings from this analysis are discussed in detail in Section 3.2 of chapter three.

The final step taken to address attrition bias involved statistical imputation. Full-information maximum likelihood (FIML) estimation was used to impute missing data in aim 1 within the context of the multiple-group latent growth curve models. FIML uses all data to estimate all parameters, including individuals who dropped out of the study (Enders, 2001; Enders & Bandalos, 2001; Heron et al., 2011). FIML implemented in Lavaan maximizes a likelihood function (i.e., the sum of *n* casewise likelihood functions) to directly estimate model parameters and standard errors using all available raw data (Enders, 2001). FIML has been shown to be accurate when data are missing at random (i.e., the drop-out rate is MAR; Peugh & Enders, 2004), as well as when data are not missing at random when additional measures are taken to explore bias (NMAR; Enders, 2011; Muthén et al., 2011). The *k*-nearest neighbors algorithm was used to impute missing data on independent variables in Aim 2 as part of data preparation for the random forest models. The k-nearest neighbor imputer (kNNImputer) from the Scikit-Learn Library (Pedregosa et al., 2011) was used with *k* equal to 5 and 'nan-euclidian' as the distance metric. The missing values of one feature are imputed by the average of those from its *k* nearest neighbors, which are calculated using the remaining features of the dataset. Both methods of imputation are sensitive and robust to different mechanisms of missingness compared to traditional methods for handling missing data (e.g., listwise deletion, mean imputation; Emmanuel et al., 2021; Enders & Bandalos, 2001; Mohammed et al., 2021).

2.4.2 Aim 1: Multiple-Group Latent Growth Analyses

The first aim was to determine whether women and men differ significantly in initial levels (i.e., intercepts) and rates of change (i.e., slopes) of recovery capital over a six-month period using a multiple-group (i.e., gender-specific) latent growth modeling approach. Multiple-group latent growth models allow for structural invariance testing, or the examination of whether latent factors are indeed equal across separate groups (e.g., women and men). Unlike traditional latent growth models, multiple-group modeling allows for latent factors (latent intercept, slope) to be estimated separately. If structural relationships differ between groups, it is thought that group membership moderates the relationship between latent factors (Sass & Schmitt, 2013), thereby justifying further gender-specific investigation. This method is fitting for this study because, as a group-specific quantitative methodological tool, it allows for the contextualization of women's and men's recovery experiences.

Before estimating the primary models (multiple-group latent growth models), a series of preliminary analyses were carried out including testing an appropriate cutoff to examine change over time. Six months was initially chosen as a developmentally appropriate period of time to observe change in recovery capital because it has previously been used in the literature as a temporal benchmark for follow-up of individuals in early recovery (Laudet & White, 2008; Polcin et al., 2010). However, the parent study's retention rate at month six (52%) was eight percentage points below the recommended threshold for substance use studies (i.e., 60%; Hansten et al., 2000). Therefore, before moving forward, a comparative invariance analysis was carried out with the depression and financial strain variables using both four months (retention = 62%) and six months (retention = 52%) as cutoffs. This analysis investigated adequacy of model fit at both time points and compared qualitative differences in the predicted intercepts and rates of change over time; results supported the six-month cutoff and are reported in section 3.3.1.

After testing the models at different time points, two additional preliminary steps were carried out prior to the estimation of the primary multiple-group latent growth models. First, a full sample unconditional model was estimated for each type of recovery capital to test whether a linear versus quadratic curve provided the best fit to the data. The second step, as recommended by Byrne et al. (1989), was to estimate the recovery capital models across women (n = 72) and men (n = 46) separately to determine whether the latent growth model adequately fit both groups. If the models did not adequately fit both groups separately then a multiple-group model may not be appropriate. Adequacy of fit for all latent growth models was examined using a series of fit indices (for a detailed description of model fit indices and corresponding cutoff values, refer to Appendix B). These preliminary analyses determined that the mindfulness and internalized stigma models did not adequately fit the men's data and were therefore not included in the primary multiple-group latent growth models. The primary analysis moved forward with the depression, perceived stress, financial strain, and social support models.

Once the preliminary tests were complete, the primary models were estimated sequentially. First, an unconditional multiple-group latent growth model was estimated for both women and men (i.e., full sample, n = 120; see Figure 2.2), freeing all 22 structural parameters (i.e., four growth factor means [one for intercept and slope of women's recovery capital, one for intercept and slope of men's recovery capital], 16 variances [12 error variances for repeated measures, six per model; four variances for growth factors, two per model], and two covariances [one between each model's growth factors]). Then, structural invariance, which refers to characteristics of the latent variables (Drapeau et al., 2010), was examined using a forward approach where constrained parameters are incrementally added to the model and compared (see Sass & Schmitt, 2013). Comparative fit indices were used to determine improvements in model fit over the baseline model (e.g., CFI, TLI, RMSEA). By incrementally exploring these elements of the model, any resultant differences in gender were more comprehensively understood and therefore interpretable.

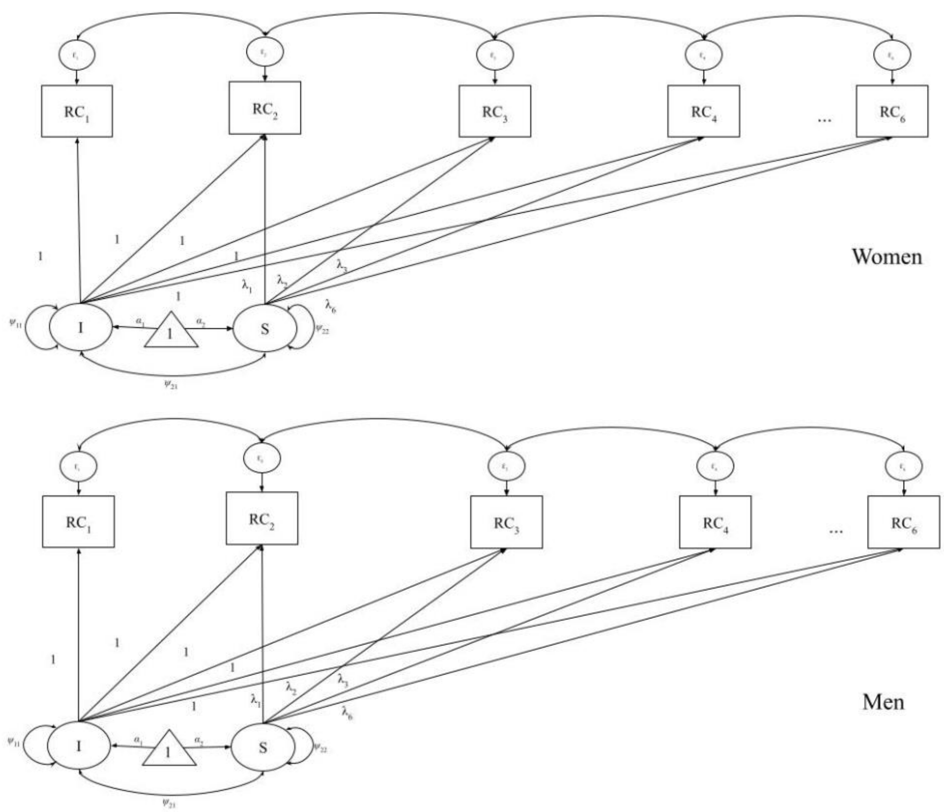

Figure 2.2 Multiple-Group Latent Growth Model for Six Repeated Measures of Recovery Capital Using Gender as Grouping Variable.

Note. RC = Recovery Capital, I = Intercept, S = Slope

2.4.3 Aim 2: Random Forest Models

The second aim was to explore variables that influence recovery using a series of random forest classification models. Random Forest (RF) is a nonparametric, ensemble learning algorithm based on decision trees that uses features (predictors) to classify (or predict) a target variable of interest (categorical outcome; Breiman, 2001). A decision tree is a binary decisions flowchart. At every step of the tree, all variables are considered to split the observations into two groups where the aim is to maximize the purity of each group with respect to the target. The ideal model is a simple tree with few splits to avoid overfitting and final nodes that are pure (i.e., all samples within a group classified into one category of outcome). Because they make binary decisions on features one at a time, tree-based methods, such as Random Forest, are also particularly suited to datasets that contain both categorical and numerical variables, such as the current dataset.

Random forest can work with small to medium data, unbalanced data (i.e., datasets where the dependent variable [target] has an uneven or skewed distribution of observations across categories), and can perform well even when independent variables (features) outnumber data objects (Pirneskoski et al., 2020; Qi, 2012). These qualities made it an appropriate fit for the current study, which featured a smaller community-based sample (n = 120) and a skewed distribution in the recovery status target variable (i.e., stable recovery class was twice as common as unstable recovery class). Furthermore, a feature importance analysis performed on a random forest provides meaningful information about the relative importance of predictors included in the analysis; such information is potentially of great use to clinical and community-based environments (e.g., sober living homes), as an understanding of which characteristics are most predictive of stable recovery for women versus men could help to inform gender-responsive policy and programming.

One of the limitations of tree methods is their tendency towards overfitting (i.e., high levels of variance) and bias. A machine learning model is biased if it systematically under or over predicts the target variable (Ramchandani, 2018). Ideally, a classifier model will achieve low bias-low variance by maximizing accuracy (i.e., fraction of correct predictions). Ensemble methods like Random Forest employ a number of techniques to reduce variance in predictions while maintaining generally low bias.

The Random Forest algorithm uses a bootstrap approach, resampling over features, observations, or both, to develop a forest of classification decision trees, which are then aggregated in terms of prediction and prediction errors. In other words, a random forest is an ensemble of decision trees where each tree works with a subset of the input variables and/or a subset of the observations. Compared to decision trees, the Random Forest decreases variance in the prediction, is more robust to overfitting (due to resampling) and produces a prediction with an associated uncertainty (i.e., a prediction that can be interpreted probabilistically).

Procedures for the random forests model are carried out in three steps. First, bootstrap samples are drawn from the original data. A portion of observations is left out of each bootstrap sample; these are referred to as “out of bag” (OOB) samples. Then, a decision tree is built using the remaining data, dubbed “training” data. From this tree, a random set of features (predictors) is drawn and used to predict the outcome. Finally, prediction estimates are made for the “out of bag” observations by aggregating the predictions from all individual trees by “majority decision” (i.e., choosing the class predicted by the majority of the trees). The OOB R^2 helps prevent overfitting by validating the model fit on unseen data (Svetnik et al., 2003). Once the outcome is predicted, the relative importance of each independent variable, commonly referred to as “feature importance”, can be calculated for each tree and its uncertainty is characterized by aggregating the results of all the trees in the forest. The importance of the input features in the two gender-based models were compared in an effort to identify gender-based differences in stable recovery.

Having selected a machine learning model, the researcher engages in a process known as feature engineering where domain knowledge is used to create predictors that are both suitable for the chosen model and insightful about the outcome. In the current study, as mentioned above, features were selected directly from the survey data (e.g., sociodemographic variables) as well as created using results from the analysis performed in Aim 1. Random Forest naturally works with features of different types. Categorical variables with more than ~3 categories are “one-hot encoded”: each of the N categories is turned into a separate, binary feature. In addition to creating features, the researcher makes model choices (e.g., how many trees to include in the model) and modifications in an effort to improve performance. This is an iterative process that, when carried out correctly, results in a model with high accuracy and minimal error.

In the current analysis, Random Forest models were built both using the full sample and separately for women and men (three models total). Results from the first aim (multiple-group latent growth models) were used as features, or predictors, together with socioeconomic, substance use, and mental health variables as input to the three random forest models. To determine the optimal parameters for each model, a parameter grid search was performed to identify the best hyperparameters separately for each of the three models, including the maximum depth of the trees (number of splits along a branch), number of trees in the forest, bootstrapping scheme, and number of features to be used in each tree.

The gender-specific models (i.e., women only, men only) addressed the primary thesis of the project (contextualization of women’s recovery experiences), and the full sample model examined the predictive importance of gender relative to other conceptually relevant variables and sociodemographic characteristics. Findings derived from these models further illustrate differences between women’s and men’s recovery for individuals residing in and transitioning from sober living homes.

Chapter 3

RESULTS

This chapter presents the results of the analyses outlined in the Methods section. The objective of the first aim was to determine whether women's and men's recovery capital differed significantly in initial levels (i.e., intercepts) and rates of change (i.e., slopes) over a six-month period. The objective of the second aim was to use gender-specific parameters from Aim 1 along with key sociodemographic, substance use, and mental health variables to explore which variables contribute to recovery status for women and for men residing in and transitioning out of sober living homes, as well as to compare similarities and differences in the relative importance of said variables across gender.

3.1 Descriptive and Exploratory Data Analysis

Before estimating the multiple-group latent growth curve models, sociodemographic and substance use predictors as well as recovery capital variables were tested for categorical and mean differences between genders (see Tables 3.1-3.3).

3.1.1 Sociodemographics

The final sample included 120 residents (61% women) residing in eight recovery homes in the state of Delaware. Residents ranged in age from 19 to 64 (M = 36.36, SD = 10.87), and, as shown in Table 3.1, most identified as white, non-Hispanic (79% white, non-Hispanic). Over 22% of the sample had either a college or graduate degree and 53% reported a monthly income of $250 or less. Table 3.1 outlines gender-specific differences across sociodemographic covariates. Notable differences were observed between women and men in terms of LGBTQ membership, monthly income, and parental status (i.e., whether or not participants reported having children).

Table 3.1 Descriptives of Sociodemographic Variables at Baseline

	Full Sample	Grouped by Gender		
Characteristic	N = 120[1]	Men, N = 47[1]	Women, N = 73[1]	*p*-value
Age*	36.4 (10.9)	36.3 (10.3)	36.4 (11.3)	0.921
Race/Ethnicity†				0.332
Black, non-Hispanic	12 (10.0%)	3 (6.4%)	9 (12.3%)	
Hispanic	9 (7.5%)	3 (6.4%)	6 (8.2%)	
Native American, non-Hispanic	1 (0.8%)	1 (2.1%)	0 (0.0%)	
Other	3 (2.5%)	0 (0.0%)	3 (4.1%)	
White, non-Hispanic	95 (79.2%)	40 (85.1%)	55 (75.3%)	
LGBTQ‡				**0.003**
0=No	96 (80.0%)	44 (93.6%)	52 (71.2%)	
1=Yes	24 (20.0%)	3 (6.4%)	21 (28.8%)	
Education Level†				0.356
Some High School or less (no diploma or GED)	18 (15.0%)	7 (14.9%)	11 (15.1%)	
Completed High School or GED	38 (31.7%)	16 (34.0%)	22 (30.1%)	
Some College or Technical School (no degree)	37 (30.8%)	11 (23.4%)	26 (35.6%)	
2-Year College Degree	10 (8.3%)	3 (6.4%)	7 (9.6%)	
4-year College Degree	14 (11.7%)	9 (19.1%)	5 (6.8%)	
Some Graduate School (no degree)	2 (1.7%)	1 (2.1%)	1 (1.4%)	
Graduate School Degree	1 (0.8%)	0 (0.0%)	1 (1.4%)	
Monthly Income†				**0.008**
$0-250	62 (53.0%)	24 (53.3%)	38 (52.8%)	
$251-500	8 (6.8%)	3 (6.7%)	5 (6.9%)	
$501-1,000	14 (12.0%)	2 (4.4%)	12 (16.7%)	
$1,001-1,500	12 (10.3%)	2 (4.4%)	10 (13.9%)	
$1,501-2,000	7 (6.0%)	3 (6.7%)	4 (5.6%)	
$2,001-more	14 (12.0%)	11 (24.4%)	3 (4.2%)	
Employed‡				0.837
0=No	83 (69.2%)	32 (68.1%)	51 (69.9%)	
1=Yes	37 (30.8%)	15 (31.9%)	22 (30.1%)	
Children‡				**< .001**
0=No	42 (35.0%)	28 (59.6%)	14 (19.2%)	

1=Yes	78 (65.0%)	19 (40.4%)	59 (80.8%)	
Relationship Status†				0.12
Dating	6 (5.0%)	3 (6.4%)	3 (4.1%)	
Divorced	17 (14.2%)	6 (12.8%)	11 (15.1%)	
In a relationship	17 (14.2%)	2 (4.3%)	15 (20.5%)	
Married	9 (7.5%)	5 (10.6%)	4 (5.5%)	
Separated	10 (8.3%)	3 (6.4%)	7 (9.6%)	
Single	60 (50.0%)	28 (59.6%)	32 (43.8%)	
Widowed	1 (0.8%)	0 (0.0%)	1 (1.4%)	
Lifetime Incarceration‡				0.507
0=No	50 (42.0%)	18 (38.3%)	32 (44.4%)	
1=Yes	69 (58.0%)	29 (61.7%)	40 (55.6%)	
Currently on Parole†				0.197
0=No	108 (90.8%)	45 (95.7%)	63 (87.5%)	
1=Yes	11 (9.2%)	2 (4.3%)	9 (12.5%)	

[1]Mean (SD); *n* (%)
*Wilcoxon Sum of Ranks test was used to compare differences between gender and non-normal continuous variables.
†Fisher's exact test was used to compare differences between gender and non-normal categorical variables.
‡Pearson's Chi-squared test was used to compare differences between gender and categorical variables with a normal distribution.

3.1.2 Substance Use and Mental Health

Table 3.2 outlines gender-specific differences across substance use and mental health covariates. At baseline, 42% of participants reported substance use within the previous three months. Primary substance of use differed significantly between women and men; further examination of the residuals (i.e., the difference between the observed counts and expected counts divided by an estimate of the standard error) of the chi-square test result indicated that alcohol (women, 14.3%; men, 32.6%) made the greatest contribution to the test's overall significance.

In terms of lifetime mental health diagnoses, 79.59% of women had three or more diagnoses, compared to 20.41% of men. Both Bipolar I[1] disorder and posttraumatic stress disorder were significantly more prevalent among women compared to men. On average, men reported significantly more days in recovery as well as more days spent in their sober living home compared to women.

---

[1] Bipolar I Disorder is characterized by manic episodes that last at least 7 days, or by manic symptoms that are so severe that the person needs immediate hospital care; and, Bipolar II Disorder is defined by depressive episodes and hypomanic episodes, but not the full-blown manic episodes that are typical of Bipolar I Disorder (National Institute of Mental Health, 2020).

Table 3.2 Descriptives of Substance Use and Mental Health Variables at Baseline

	Full Sample	Grouped by Gender		
Characteristic	N = 120[1]	Men, N = 47[1]	Women, N = 73[1]	*p*-value
Number of Lifetime Overdoses*	1.5 (2.7)	1.4 (2.7)	1.6 (2.6)	.257
Primary Substance of Use†				**.042**
Alcohol	25 (21.6%)	15 (32.6%)	10 (14.3%)	
Coke	20 (17.2%)	4 (8.7%)	16 (22.9%)	
Opioids	60 (51.7%)	24 (52.2%)	36 (51.4%)	
Other	11 (9.5%)	3 (6.5%)	8 (11.4%)	
Addiction Severity*	25.6 (5.8)	26.1 (4.3)	25.3 (6.6)	.676
Number of Lifetime Mental Health Diagnoses†				**.030**
0	16 (13.7%)	9 (20.5%)	7 (9.6%)	
1	24 (20.5%)	14 (31.8%)	10 (13.7%)	
2	28 (23.9%)	11 (25.0%)	17 (23.3%)	
3	28 (23.9%)	8 (18.2%)	20 (27.4%)	
4	10 (8.5%)	2 (4.5%)	8 (11.0%)	
5	4 (3.4%)	0 (0.0%)	4 (5.5%)	
6	5 (4.3%)	0 (0.0%)	5 (6.8%)	
7	2 (1.7%)	0 (0.0%)	2 (2.7%)	
Mean Days in Sober Living Home*	89.8 (163.9)	124.7 (155.0)	67.3 (166.5)	**.036**
Mean Days in Recovery*	558.5 (1,186.2)	815.8 (1,463.7)	385.8 (928.4)	**.005**
Lifetime Relapse				**.009**
0 Relapses	13 (11.3%)	10 (21.28%)	3 (23.08%)	
1-3 Relapses	44 (38.26%)	13 (27.66%)	31 (70.45%)	
3+ Relapses	58 (50.43%)	24 (51.06%)	34 (58.62%)	
SLH Org 1 vs. SLH Org 2†				**< .001**
SLH Org 2	57 (47.5%)	8 (17.0%)	49 (67.1%)	
SLH Org 1	63 (52.5%)	39 (83.0%)	24 (32.9%)	
Diagnosed with Bipolar1†				**< .001**
No	86 (73.5%)	40 (90.9%)	46 (63.0%)	
Yes	31 (26.5%)	4 (9.1%)	27 (37.0%)	
Diagnosed with Depression‡				.520
No	62 (53.0%)	25 (56.8%)	37 (50.7%)	
Yes	55 (47.0%)	19 (43.2%)	36 (49.3%)	
Diagnosed with PTSD†				**< .001**
No	73 (62.4%)	36 (81.8%)	37 (50.7%)	
Yes	44 (37.6%)	8 (18.2%)	36 (49.3%)	

Diagnosed with Anxiety Disorder‡				.278
No	51 (43.6%)	22 (50.0%)	29 (39.7%)	
Yes	66 (56.4%)	22 (50.0%)	44 (60.3%)	
Recovery Status (across six months) ‡				
Stable	82 (68.3%)	35 (74.5%)	47 (64.4%)	.246
Unstable	38 (31.7%)	12 (25.5%)	26 (35.6%)	

Note. SLH Org = Sober living home organization. Column percentages are reported.
[1]Mean (SD); *n* (%)
*Wilcoxon Sum of Ranks test was used to compare differences between gender and non-normal continuous variables.
†Fisher's exact test was used to compare differences between gender and non-normal categorical variables.
‡Pearson's Chi-squared test was used to compare differences between gender and categorical variables with a normal distribution.

3.1.3 Recovery Capital

Raw means and standard deviations for recovery capital variables, as well as gender-specific differences are reported in Table 3.3. Significant differences between women's and men's recovery capital were observed at one or more time points for depression, internalized stigma, financial strain, and social support.

Table 3.3 Descriptives of Recovery Capital Variables

Characteristic	Full Sample N = 120[1]	Grouped by Gender Men, N = 47[1]	Women, N = 73[1]	p-value[2]	Missing[3]
Depression					
Baseline	10.6 (6.3)	9.3 (6.2)	11.5 (6.3)	.081	14 (12%)
Month 2	10.6 (7.1)	10.6 (5.9)	10.7 (8.0)	.923	30 (25%)
Month 3	9.3 (6.6)	8.9 (6.6)	9.7 (6.6)	.627	45 (38%)
Month 4	9.8 (7.6)	9.5 (7.5)	10.1 (7.9)	.741	49 (41%)
Month 5	**10.4 (8.0)**	**7.9 (6.6)**	**12.4 (8.4)**	**.021**	**56 (47%)**
Month 6	8.8 (6.7)	7.1 (5.8)	10.0 (7.1)	.101	63 (52%)
Mindfulness					
Baseline	2.8 (1.2)	3.0 (1.2)	2.7 (1.2)	.314	3 (2.5%)
Month 2	2.8 (1.0)	2.9 (1.0)	2.8 (1.0)	.371	26 (22%)
Month 3	2.9 (1.1)	2.9 (1.1)	2.9 (1.2)	.889	40 (33%)
Month 4	3.0 (1.0)	3.0 (1.1)	3.0 (0.9)	.972	46 (38%)
Month 5	3.2 (1.1)	3.3 (1.2)	3.0 (1.1)	.315	54 (45%)
Month 6	3.2 (1.1)	3.3 (0.9)	3.1 (1.2)	.404	59 (49%)
Perceived Stress					
Baseline	7.1 (3.0)	6.6 (2.5)	7.4 (3.2)	.115	6 (5.0%)
Month 2	6.6 (2.7)	6.5 (2.4)	6.8 (3.0)	.651	26 (22%)
Month 3	6.5 (3.0)	6.0 (2.7)	7.0 (3.3)	.121	40 (33%)
Month 4	5.9 (3.2)	5.9 (2.9)	5.8 (3.5)	.864	46 (38%)
Month 5	6.0 (3.2)	5.2 (2.8)	6.6 (3.3)	.062	56 (47%)
Month 6	5.8 (2.9)	5.6 (2.9)	5.9 (2.9)	.665	58 (48%)
Internalized Stigma					
Baseline	2.5 (1.1)	2.4 (1.1)	2.5 (1.2)	.784	5 (4.2%)
Month 2	1.9 (1.1)	1.8 (1.0)	1.9 (1.1)	.544	25 (21%)
Month 3	1.7 (1.1)	1.5 (0.9)	1.8 (1.2)	.272	38 (32%)
Month 4	1.5 (1.0)	1.3 (1.0)	1.7 (1.1)	.112	46 (38%)
Month 5	**1.8 (1.1)**	**1.4 (0.7)**	**2.1 (1.3)**	**.003**	**54 (45%)**
Month 6	**1.6 (1.2)**	**1.1 (0.8)**	**1.9 (1.3)**	**.002**	**58 (48%)**
Financial Strain					
Baseline	**14.2 (6.8)**	**12.0 (5.9)**	**15.6 (7.0)**	**.004**	**6 (5.0%)**
Month 2	**14.4 (6.8)**	**12.8 (6.4)**	**15.8 (6.9)**	**.031**	**28 (23%)**
Month 3	**12.7 (6.2)**	**11.2 (5.6)**	**14.1 (6.5)**	**.037**	**43 (36%)**
Month 4	**12.9 (6.8)**	**11.0 (5.9)**	**14.6 (7.2)**	**.023**	**48 (40%)**
Month 5	**12.5 (6.4)**	**9.7 (4.9)**	**14.6 (6.7)**	**.001**	**56 (47%)**
Month 6	**12.2 (6.7)**	**9.6 (4.1)**	**14.3 (7.6)**	**.003**	**59 (49%)**

Social Support					
Baseline	**2.7 (1.0)**	**3.0 (0.7)**	**2.6 (1.2)**	**.026**	**3 (2.5%)**
Month 2	2.9 (1.0)	2.9 (0.9)	3.0 (1.0)	.582	25 (21%)
Month 3	3.1 (0.9)	3.0 (0.8)	3.1 (0.9)	.667	40 (33%)
Month 4	3.0 (1.0)	3.0 (0.9)	3.0 (1.0)	.956	46 (38%)
Month 5	3.1 (0.9)	3.2 (0.8)	3.0 (0.9)	.396	54 (45%)
Month 6	3.1 (1.0)	3.1 (1.0)	3.1 (1.1)	.967	60 (50%)

[1]Mean (SD)
[2]Welch Two Sample t-test
[3]N Missing (% Missing)

3.2 Attrition Analyses

Bivariate analyses were used to examine differences between individuals with complete versus incomplete data across all patterns of missingness using sociodemographic, substance use, and mental health predictors at baseline. Results indicated that race/ethnicity (i.e., Other versus White, non-Hispanic), LGBTQ status, education (≤ High School Degree versus Some College or more), primary substance of use (i.e., opioids versus other), number of days in sober living home at baseline, sober living home organization affiliation, and whether or not participants moved out during the study predicted significant differences between individuals with complete versus incomplete data (see Table 3.4).

Table 3.4 Bivariate Analyses Examining Differences in Complete versus Incomplete Data using Sociodemographic, Substance Use, and Mental Health Predictors at Baseline

Characteristic	Missing, N = 89[1]	Complete, N = 31[1]	*p*-value[2]
Gender			0.427
Men	33 (37.1%)	14 (45.2%)	
Women	56 (62.9%)	17 (54.8%)	
Age	35.3 (10.0)	39.3 (12.8)	0.121
Race/Ethnicity			**0.022**
Other	23 (25.8%)	2 (6.5%)	
White, non-Hispanic	66 (74.2%)	29 (93.5%)	
LGBTQ			**0.029**
No	67 (75.3%)	29 (93.5%)	
Yes	22 (24.7%)	2 (6.5%)	
Education			**0.002**
≥ Some College or More	40 (44.9%)	24 (77.4%)	
≤ High School Degree	49 (55.1%)	7 (22.6%)	
Income			0.647
No	36 (40.4%)	14 (45.2%)	
Yes	53 (59.6%)	17 (54.8%)	
Employed			0.801
No	61 (68.5%)	22 (71.0%)	
Yes	28 (31.5%)	9 (29.0%)	
Children			0.07
No	27 (30.3%)	15 (48.4%)	
Yes	62 (69.7%)	16 (51.6%)	
Relationship Status			0.835
No	44 (49.4%)	16 (51.6%)	
Yes	45 (50.6%)	15 (48.4%)	
Lifetime Incarceration			0.208
No	34 (38.6%)	16 (51.6%)	
Yes	54 (61.4%)	15 (48.4%)	
Currently on Parole			>.999
No	81 (91.0%)	28 (90.3%)	
Yes	8 (9.0%)	3 (9.7%)	
Primary Substance			**0.002**
Other	37 (41.6%)	23 (74.2%)	
Opioids	52 (58.4%)	8 (25.8%)	
Addiction Severity			0.54
Mild (0-20)	14 (15.7%)	5 (16.1%)	
Moderate (21-25)	20 (22.5%)	4 (12.9%)	

Severe (26-30)	55 (61.8%)	22 (71.0%)	
Number of Mental Health Disorder Diagnoses			0.575
0	10 (11.6%)	6 (19.4%)	
1	20 (23.3%)	4 (12.9%)	
2	21 (24.4%)	7 (22.6%)	
3	18 (20.9%)	10 (32.3%)	
4	8 (9.3%)	2 (6.5%)	
5	4 (4.7%)	0 (0.0%)	
6	3 (3.5%)	2 (6.5%)	
7+	2 (2.3%)	0 (0.0%)	
Number of Days in Recovery	480.2 (1,010.3)	775.7 (1,575.5)	0.279
Number of Days in Sober Living Home	**84.0 (174.7)**	**106.3 (128.8)**	**0.007**
Number of Lifetime Relapses			0.43
0 Relapses	8 (9.5%)	5 (16.1%)	
1-3 Relapses	31 (36.9%)	13 (41.9%)	
3+ Relapses	45 (53.6%)	13 (41.9%)	
Sober Living Home Organization			**0.005**
Organization 2	49 (55.1%)	8 (25.8%)	
Organization 1	40 (44.9%)	23 (74.2%)	
Moved Out of Home During Study			**< .001**
No	31 (34.8%)	23 (74.2%)	
Yes	58 (65.2%)	8 (25.8%)	
Moved Out of Home in < 30 days			0.112
No	75 (84.3%)	30 (96.8%)	
Yes	14 (15.7%)	1 (3.2%)	
Experienced Relapse During Study			0.717
No	69 (77.5%)	25 (80.6%)	
Yes	20 (22.5%)	6 (19.4%)	

Note. Results in bold indicate significant differences between gender.
[1] Mean (SD)
[2] Pearson's Chi-squared test; Wilcoxon rank sum test; Fisher's exact test

A logit analysis was used to predict 100% complete vs incomplete data. A logit model takes into account the effect of a given predictor while holding all other variables constant and provides additional information about the direction of a relationship. The parameter estimates of the logit model represent the logarithm of the odds of completion across 100% of the variables and time points based on various independent variables at baseline (see Table 3.5). The odds of completion can be calculated by exponentiating the logarithm. The logit analysis results indicated that being White, non-Hispanic predicted higher odds of completing 100% of the surveys, whereas LGBTQ status, lower education (≤ High School Degree versus Some College or more), opioid use, and moving out of the sober living home during the course of the study predicted lower odds of completing all surveys.

Table 3.5 Logit Analysis of Independent Variables at Baseline Comparing Completion versus Attrition Groups

Baseline Variable	**ß**	**Exp(ß)**	**Wald**
Women	0.91	2.48	0.85
Age	0.02	1.02	0.38
White	**2.72***	**15.21**	**2.43**
LGBTQ	**-2.33***	**0.1**	**-2.01**
Education: <= High School Degree	**-1.88***	**0.15**	**-2.08**
Income: ≤ $500/week	-1.1	0.33	-1.21
Employed (Full-time/Part-time)	-0.75	0.47	-0.63
Parental Status (Has Children)	-0.55	0.58	-0.63
Single	-0.76	0.47	-0.97
Previously in Prison/Jail	-0.97	0.38	-0.98
Currently on Parole	2.28	9.74	1.37
Primary Substance: Opioids	**-2.66****	**0.07**	**-2.8**
Addiction Severity			
Mild (0-20)	Ref	Ref	Ref
Moderate (21-25)	-0.84	0.43	-0.66
Severe (26-30)	1.01	2.75	0.89
Number of mental health disorders (centered)	0.26	1.29	0.79
Days in Recovery (centered)	0	1	1.58
Days in Sober living home (centered)	0	1	-0.58
Number of Relapses Experienced (Lifetime)			
0	Ref	Ref	Ref
1-3	0.42	1.51	0.34
3+	-1.44	0.24	-1.15
Sober Living Home Organization	0.58	1.79	0.58
Moved Out of Home During Study	**-1.91***	**0.15**	**-2.12**
Moved Out of Home (< 30 days)	-2.73	0.07	-1.34
Experienced Relapse During Study	-1.42	0.24	-1.45

Note. Model chi-square = 63.89, *df*= 23, $^{*}p < .05$, $^{**}p < .01$. The parameter estimates represent the log of the odds of 100% participation across six waves.

3.3 Multiple-Group Latent Growth Curve Analyses

The primary objective of Aim 1 was to determine whether initial levels and rates of change in recovery capital differed significantly between women and men over a six-month period using a series of multiple-group latent growth models.

3.3.1 Preliminary Analyses

Before moving forward with the multiple-group latent growth curve analysis at six months, a comparative invariance analysis was carried out with two types of recovery capital (i.e., depression, financial strain) using both four months (retention = 62%) and six months (retention = 52%) as cutoffs (see Tables 3.6-3.7). This analysis helped to determine whether the use of six time points was appropriate for the current data despite retention at month six being below the recommended threshold. This comparative invariance analysis investigated adequacy of model fit as well as qualitative differences in the predicted intercepts and rates of change over time.

Table 3.6 Comparison of Depression Multiple-Group Latent Growth Models with Four- and Six Time Points

Fit Indices: Months 1 - 4	Parameters	χ^2	*df*	*p*-value	CFI	TLI	RMSEA	SRMR
Model 1: Means, Variances, Covariances, Residual Variances	12	18.38	16	.302	0.971	0.979	0.05	12
Model 2: Means, Variances, Residual Variances	**11**	**18.42**	**17**	**.363**	**0.983**	**0.988**	**0.038**	**11**
Model 3: Means, Variances, Covariances	11	20.15	17	.267	0.962	0.973	0.056	11
Model 4: Means and Variances	10	20.35	18	.313	0.972	0.981	0.047	10
Model 5: Variances Only	8	95.60	20	< .001	0.094	0.456	0.254	8
Model 6: Means only	8	21.68	20	.358	0.98	0.988	0.038	8
Fit Indices: Months 1 - 6	Parameters	χ^2	*df*	*p*-value	CFI	TLI	RMSEA	SRMR
Model 1: Means, Variances, Covariances, Residual Variances	12	60.26	42	.034	0.902	0.93	0.086	12
Model 2: Means, Variances, Residual Variances	**11**	**60.26**	**43**	**.042**	**0.907**	**0.935**	**0.083**	**11**
Model 3: Means, Variances, Covariances	11	69.39	43	.007	0.858	0.901	0.102	11
Model 4: Means and Variances	10	70.30	44	.007	0.859	0.904	0.101	10
Model 5: Variances Only	8	145.42	46	< .001	0.467	0.652	0.192	8
Model 6: Means only	8	72.3	46	.008	0.859	0.908	0.099	8

Note. n = 117. Across both models with four- and with six time points, Model 2 (bolded) demonstrated the best fit. Missing data were estimated using full-information maximum likelihood (FIML). $\chi 2$ = chi-square goodness of fit test, CFI = Comparative Fit Index (Bentler, 1990), TLI = Tucker-Lewis Index (Tucker & Lewis, 1973), RMSEA = root mean square error of approximation (Steiger, 1990), and SRMR = standardized root mean-squared residual (Bentler, 1995). The best fitting model (bold) was chosen based on overall fit using the following cutoff recommendations: RMSEA < .06, p < .05; CFI/TLI ≥ .95; SRMR < .08.

Table 3.7 Comparison of Financial Strain Multiple-Group Latent Growth Models with Four- and Six Time Points

Fit Indices: Months 1 - 4	Parameters	χ^2	*df*	*p*-value	CFI	TLI	RMSEA	SRMR
Model 1: Means, Variances, Covariances, Residual Variances	12	16.67	16	0.407	0.995	0.996	0.03	0.09
Model 2: Means, Variances, Residual Variances	**11**	**16.72**	**17**	**0.474**	**1**	**1.002**	**0**	**0.09**
Model 3: Means, Variances, Covariances	11	16.87	17	0.463	1	1.001	0	0.09
Model 4: Means and Variances	10	17.08	18	0.518	1	1.005	0	0.09
Model 5: Variances Only	8	106.11	20	0	0.321	0.592	0.27	1.49
Model 6: Means only	8	18.49	20	0.555	1	1.007	0	0.13
Fit Indices: Months 1 - 6	Parameters	χ^2	*df*	*p*-value	CFI	TLI	RMSEA	SRMR
Model 1: Means, Variances, Covariances, Residual Variances	12	63.31	42	0.018	0.912	0.937	0.09	0.16
Model 2: Means, Variances, Residual Variances	**11**	**63.37**	**43**	**0.023**	**0.916**	**0.941**	**0.09**	**0.16**
Model 3: Means, Variances, Covariances	11	64.6	43	0.018	0.91	0.938	0.09	0.17
Model 4: Means and Variances	10	64.6	44	0.023	0.915	0.942	0.09	0.17
Model 5: Variances Only	8	161.64	46	0	0.521	0.687	0.21	1.44
Model 6: Means only	8	71.41	46	0.01	0.895	0.931	0.1	0.3

Note. n = 119. Across both models with four- and with six time points, Model 2 (bolded) demonstrated the best fit. Missing data were estimated using full-information maximum likelihood (FIML). $\chi 2$ = chi-square goodness of fit test, CFI = Comparative Fit Index (Bentler, 1990), TLI = Tucker-Lewis Index (Tucker & Lewis, 1973), RMSEA = root mean square error of approximation (Steiger, 1990), and SRMR = standardized root mean-squared residual (Bentler, 1995). The best fitting model (bold) was chosen based on overall fit using the following cutoff recommendations: RMSEA < .06, p < .05; CFI/TLI ≥ .95; SRMR < .08.

Both sets of models demonstrated acceptable fit. While the overall model fit proved to be better for the four-month models compared to the six-month models (see Table 3.7), the six-month models provided additional information about the overall trend in recovery capital over time, a primary aim of the study. Indeed, in several cases, the recovery capital trajectories were significantly different when measured over the six-month period, compared to the four-month. The following figures show the results for the multiple-group latent growth models, visualized separately for women and men.

Depression: Figure 3.1 shows differences in women's and men's mean depression at baseline as well as its rate of change over four months versus six months. Women's depression in the four-month model (μ intercept = 11.44, $p < .001$, μ slope = -0.35, $p = .367$) appears to be decreasing but with a non-significant slope; by month 6 (μ intercept = 11.13, $p < .001$; μ slope = 0.09, $p = .703$), this downward trend flattens and its significance decreases. Across both the four-month and six-month models, the depression slope did not reach statistical significance, which implies that a large variance exists in the individual depression trajectories. While non-significant, the change in women's slope from month four to month six suggests that the addition of two months may contribute meaningful information about women's depression over time.

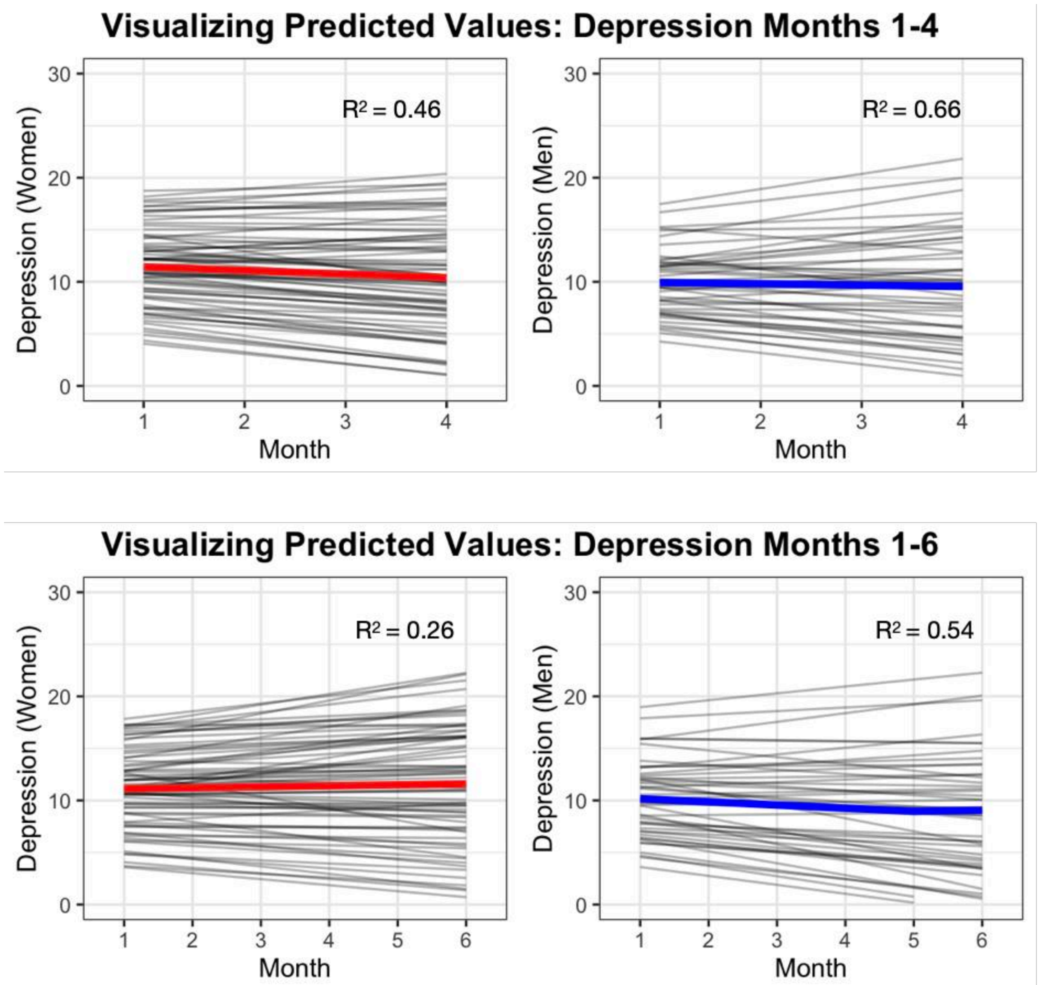

Figure 3.1 Time Series Plots Comparing Predicted Depression Trajectories for Women and Men over Four Months and Six Months

Note. R^2 was calculated as a measure of overall goodness of fit using the sum of squared residuals and the total sum of squares.

Financial strain: Figure 3.2 shows differences in women's and men's mean financial strain at baseline as well as its rate of change over four months and six months; men's financial strain in the four-month model (μ intercept = 12.47, $p < .001$,

μ slope = -0.44, $p = .182$) shows a slight decrease that is not significant. However, the addition of two months reveals a steeper and significant decrease over time in men's financial strain (μ intercept = 12.63, $p < .001$; μ slope = -0.61, $p < .001$), differentiating them from women's financial strain in terms of both initial status (μ intercept = 15.59, $p < .001$) and rate of change over time (μ slope = -0.27, $p = .188$).

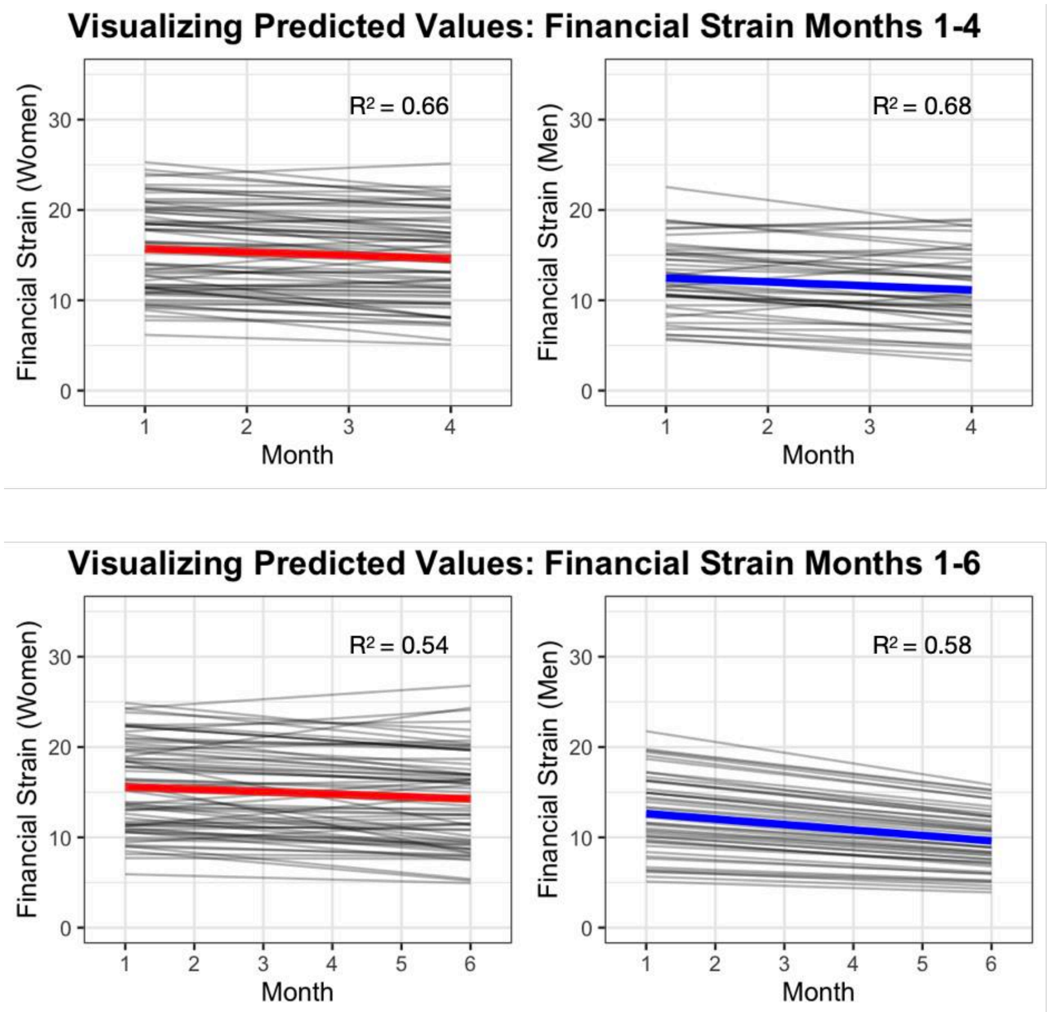

Figure 3.2 Time Series Plots Comparing Predicted Financial Strain Trajectories for Women and Men over Four Months and Six Months.

Note. R^2 was calculated as a measure of overall goodness of fit using the sum of squared residuals and the total sum of squares.

These results, in general, indicate that observing change over a six-month period reveals additional information about women's and men's recovery capital over time. Limiting the time series in Aim 1 to four months would increase the amount of complete data by 10% but decrease the time series data by 33%. Given the meaningful differences in rate of change observed by month 6 compared to month 4 across two types of recovery capital, it was determined that the tradeoff of additional complete data (i.e., 10% increase) in exchange for fewer time series (i.e., 33% decrease) was not worthwhile. Thus, six months was chosen as the analytic period of observation, and the remaining preliminary analyses were carried out over this timeline. Findings were interpreted in light of the attrition-related limitation.

The second preliminary step was to estimate a full sample unconditional model for each type of recovery capital to test whether a linear versus quadratic curve provided the best fit to the data. The linear growth curve models estimated normally (i.e., without encountering problems) for each type of recovery capital, however, the quadratic models produced multiple estimation problems including negative variance estimates and non-converged solutions indicating that the linear growth curves provided a better representation of the data. In other words, change over time for these data was constant and occurred in one direction (i.e., straight line).

The third preliminary step, as recommended by Byrne et al. (1989), was to estimate the recovery capital models across women and men separately to determine whether the proposed models adequately fit both groups. Adequacy of fit was examined using the previously discussed fit indices (see Table 3.8). Across both genders, the *mindfulness* model demonstrated poor fit as measured by a chi-square (women: χ^2 (21) = 52.12, $p < .001$; men: χ^2 (21) = 44.57, $p = .002$), RMSEA (women: RMSEA = 0.142; men: RMSEA = 0.155), and SRMR (women: SRMR = 0.17; men: SRMR = 0.21) values. The *internalized stigma* model also demonstrated less than adequate fit. In addition to convergence problems (i.e., negative variance), poor fit was observed across multiple indices including the chi square (women: χ^2 (21) = 44.26, p = .002; men: χ^2 (21) = 39.91, p = .008), RMSEA (women: RMSEA = 0.123; men: RMSEA = 0.138), and SRMR (women: SRMR = 0.17; men: SRMR = 0.21). As a result, neither the mindfulness model nor the internalized stigma model was retained for further analysis in the multiple-group framework.

Table 3.8 Fit Indices for Gender-Specific Latent Growth Models

	N	χ^2	*df*	*p*-value	CFI	TLI	RMSEA	SRMR
Depression								
Women	70	23.922	21	0.297	0.96††	0.971	0.045‡‡	0.14
Men	47	36.335	21	0.02	0.866†	0.904	0.125‡	0.17
Mindfulness								
Women	73	52.115	21	0	0.798	0.855	0.142	0.17
Men	47	44.57	21	0.002	0.864†	0.903	0.155	0.21
Perceived Stress								
Women	73	25.007	21	0.247	0.953††	0.966	0.051‡‡	0.152
Men	46	33.574	21	0.04	0.896†	0.926	0.114‡	0.144
Internalized Stigma								
Women	73	44.258	21	0.002	0.805	0.861	0.123	0.106
Men	47	39.912	21	0.008	0.796	0.854	0.138	0.204
Financial Strain								
Women	73	21.802	21	0.411	0.994††	0.996	0.023‡‡	0.091
Men	46	41.504	21	0.005	0.805†	0.86	0.146‡	0.233
Social Support								
Women	73	29.73	21	0.098	0.883††	0.916	0.075‡‡	0.182
Men	47	35.31	21	0.026	0.87†	0.907	0.12‡	0.232

Note. Missing data were estimated using full-information maximum likelihood (FIML). $\chi2$ = chi-square goodness of fit test, CFI = Comparative Fit Index (Bentler, 1990), TLI = Tucker-Lewis Index (Tucker & Lewis, 1973), RMSEA = root mean square error of approximation (Steiger, 1990), and SRMR = standardized root mean-squared residual (Bentler, 1995). Adjusted CFI and RMSEA fit indices were determined using equivalence testing (Marcoulides & Yuan, 2017; Yuan et al., 2016) via the Dynamic Model Fit app, version 1.1.0 (McNeish & Wolf, 2020).

†† Indicates the women's model satisfied excellent fit for the adjusted CFI value.

† Indicates the men's model satisfied excellent fit for the adjusted CFI value.

‡‡ Indicates the women's model satisfied excellent fit for the adjusted RMSEA value.

‡ Indicates the men's model satisfied excellent fit for the adjusted RMSEA value

Next, the primary analyses for Aim 1 were carried out. The retained recovery capital models for the primary analyses included: *depression, perceived stress, financial strain,* and *social support*. Structural invariance, or the equality of latent factor means, variances, residual variances, and covariances across gender, was tested using a series of multiple-group latent growth models. The joint model building procedure involved estimating a fully unconstrained model with 18 freely estimated parameters followed by progressively more restricted models where parameters (e.g., latent factor means, variances, residual variances, and covariances) were sequentially constrained and estimated jointly for women and men. Improvement in model fit was examined using a variety of fit indices. Finally, an optimal model was chosen for each recovery capital type based on overall improvement in model fit. Each of the recovery capital models included six indicators (corresponding to six waves for each variable measured), which were used to define two unobserved latent growth factors (i.e., intercept and slope). The fit of each model within recovery capital as well as the parameter estimates of each optimal model are discussed below (see Tables 3.9-3.12).

3.3.2 Depression

The optimal model for the *depression* data proved to be the model with means, variances, and residual variances freely estimated across gender (χ^2 (43) = 60.26, p = .042; CFI = 0.907, RMSEA = 0.083, SRMR = 0.152; Model 2 in Table 3.9). The means of the intercepts for depression were significant for both women (μ = 11.13, p < .001) and men (μ = 10.16, p < .001). Although women's average depression symptoms at baseline (i.e., intercept mean) were, on average, 0.97 points higher than that of men, a follow-up independent samples t-test determined that this difference was not statistically significant. The variances of the latent intercept were also significant for both women (ψ = 20.88, p = .001) and men (ψ = 17.30, p = .002), indicating that both women and men showed meaningful individual variability around their respective group averages at baseline.

Table 3.9 Nested Comparison of Multiple Group Latent Growth Model Fit Indices for Depression

Fit Indices	Parameters	χ^2	*df*	*p*-value	CFI	TLI	RMSEA	SRMR
Model 1: Means, Variances, Covariances, Residual Variances	12	60.26	42	0.034	0.902	0.93	0.086	0.152
Model 2: Means, Variances, Residual Variances	**11**	**60.26**	**43**	**0.042**	**0.907**	**0.935**	**0.083**	**0.152**
Model 3: Means, Variances, Covariances	11	69.39	43	0.007	0.858	0.901	0.102	0.165
Model 4: Means and Variances	10	70.3	44	0.007	0.859	0.904	0.101	0.164
Model 5: Variances Only	8	145.42	46	0	0.467	0.652	0.192	0.883
Model 6: Means only	8	72.3	46	0.008	0.859	0.908	0.099	0.189

Note. n = 117. Model 2 (bolded) demonstrated the best fit and was retained for further analysis in aim 2. Missing data were estimated using full-information maximum likelihood (FIML). $\chi2$ = chi-square goodness of fit test, CFI = Comparative Fit Index (Bentler, 1990), TLI = Tucker-Lewis Index (Tucker & Lewis, 1973), RMSEA = root mean square error of approximation (Steiger, 1990), and SRMR = standardized root mean-squared residual (Bentler, 1995). The best fitting model (bold) was chosen based on overall fit using the following cutoff recommendations: RMSEA < .06, p < .05; CFI/TLI ≥ .95; SRMR < .08. Once chosen, adjusted CFI and RMSEA values were determined using equivalence testing (Marcoulides & Yuan, 2017; Yuan et al., 2016) via the Dynamic Model Fit app, version 1.1.0 (McNeish & Wolf, 2020).
† Indicates excellent fit for the adjusted CFI value (≥ 0.914).
‡ Indicates excellent fit for the adjusted RMSEA value (≤ 0.063).

Neither the depression slope means, nor the slope variances were significant for either gender, indicating that no meaningful change was observed over time for either group in terms of overall average (i.e., slope mean) or individual variability (slope variance). The residual variances, however, were significant for both genders, indicating that unexplained individual variability in depression was observed across both groups (see Table 3.9). Finally, the relationship between the intercept and slope factors (i.e., covariance) was not significant (see Figure 3.3).

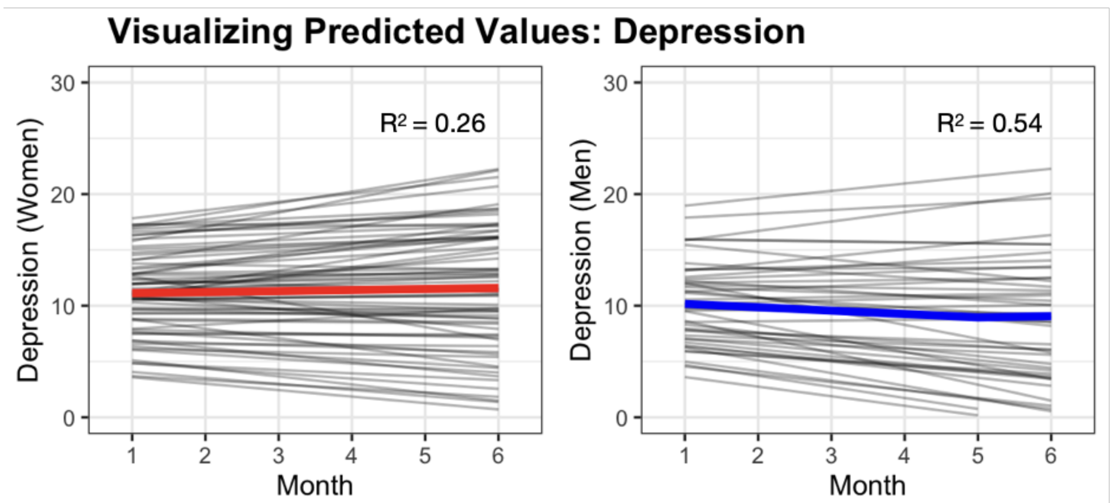

Figure 3.3 Predicted Depression Trajectories for Women and Men over Six-Month Period.

Note. R^2 was calculated as a measure of overall goodness of fit using the sum of squared residuals and the total sum of squares.

3.3.3 Perceived Stress

Similar to the depression model, the optimal model for the *perceived stress* data proved to be the model with means, variances, and residual variances freely estimated across gender (χ^2 (43) = 58.79, p = .055; CFI = 0.923, RMSEA = 0.079, SRMR = 0.149; Model 2 in Table 3.10). The means of the intercepts for perceived stress were significant for both women (μ = 7.22, p < .001) and men (μ = 6.54, p < .001). Women's average perceived stress at baseline (i.e., intercept mean) was 0.68 points higher than that of men; a follow-up independent samples *t*-test determined that this difference was statistically significant $t(117) = -2.10$, p = .038, d = -0.39 with a small effect size (d = -0.39; Cohen, 1988). The variances of the intercepts were also significant for both women (ψ = 4.73, p < .001) and men (ψ = 3.36, p < .001), indicating that both women and men demonstrated meaningful within-group variability around their respective group averages at baseline (i.e., intercept variances).

Table 3.10 Nested Comparison of Multiple Group Latent Growth Models for Perceived Stress

Perceived Stress	Parameters	χ^2	df	p-value	CFI	TLI	RMSEA	SRMR
Model 1: Means, Variances, Covariances, Residual Variances	12	58.58	42	0.046	0.919	0.942	0.081	0.149
Model 2: Means, Variances, Residual Variances	**11**	**58.79**	**43**	**.055***	**0.923†**	**0.946**	**0.079**	**0.149**
Model 3: Means, Variances, Covariances	11	75.25	43	0.002	0.843	0.891	0.112	0.176
Model 4: Means and Variances	10	75.59	44	0.002	0.846	0.895	0.11	0.18
Model 5: Variances Only	8	187	46	< .001	0.315	0.553	0.227	1.905
Model 6: Means only	8	77.35	46	0.003	0.848	0.901	0.107	0.213

Note. n = 119. Model 2 (bolded) demonstrated the best fit and was retained for further analysis in aim 2. Missing data were estimated using full-information maximum likelihood (FIML). χ2 = chi-square goodness of fit test, CFI = Comparative Fit Index (Bentler, 1990), TLI = Tucker-Lewis Index (Tucker & Lewis, 1973), RMSEA = root mean square error of approximation (Steiger, 1990), and SRMR = standardized root mean-squared residual (Bentler, 1995). The best fitting model (bold) was chosen based on overall fit using the following cutoff recommendations: RMSEA < .06, p < .05; CFI/TLI ≥ .95; SRMR < .08. Once chosen, adjusted CFI and RMSEA values were determined using equivalence testing (Marcoulides & Yuan, 2017; Yuan et al., 2016) via the Dynamic Model Fit app, version 1.1.0 (McNeish & Wolf, 2020).

* Indicates null model should be rejected.

† Indicates excellent fit for the adjusted CFI value (≥ 0.915).

‡ Indicates excellent fit for the adjusted RMSEA value (≤ .063)

While men’s perceived stress demonstrated no meaningful change over time, the mean of the women’s slope factor ($\mu = -0.21$, $p = .033$) was significant, indicating that, on average, women’s perceived stress *decreased* over the six-month period. Inversely, the variance of the men’s slope factor was significant, demonstrating meaningful individual variability in change of men’s perceived stress ($\psi = 0.2$, $p = .016$); no change was observed in the women’s group. The residual variances were significant for both genders, indicating that unexplained individual variability was observed across both groups (see Table 3.10). Finally, the relationship between the intercept and slope factors (i.e., covariance) was not significant (see Figure 3.4).

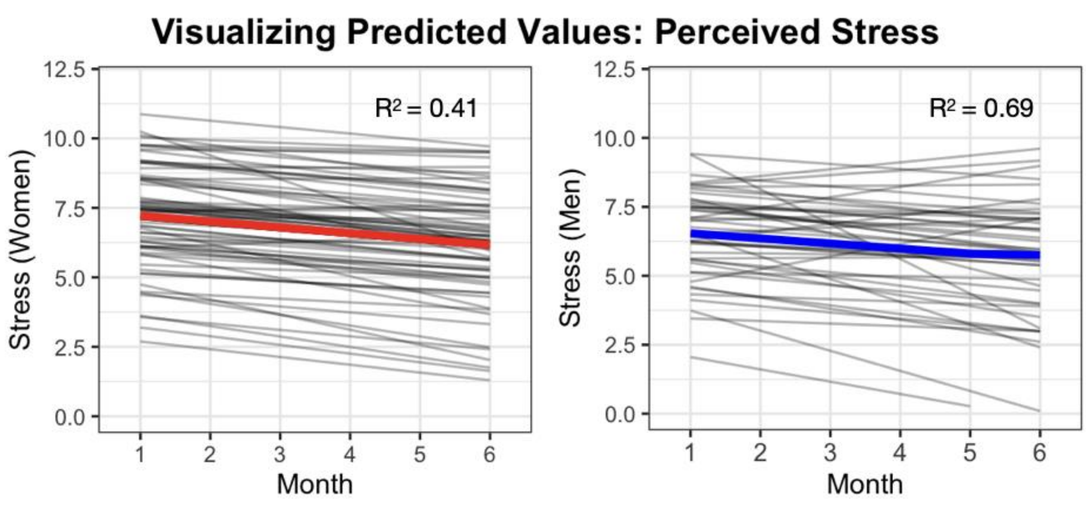

Figure 3.4 Predicted Perceived Stress Trajectories for Women and Men over Six-Month Period

Note. R^2 was calculated as a measure of overall goodness of fit using the sum of squared residuals and the total sum of squares.

3.3.4 Financial Strain

The same model proved to be the optimal model for *financial strain*, wherein factor means, variances, and residual variances were freely estimated across gender (χ^2 (43) = 63.37, p = .023; CFI = 0.916, RMSEA = 0.089, SRMR = 0.163; Model 2, Table 3.11). The means of the intercepts for financial strain were significant for both women (μ = 15.59, p < .001) and men (μ = 12.63, p < .001). Women's average financial strain at baseline (i.e., intercept mean) was 2.96 points higher than that of men; a follow-up independent samples t-test determined that this difference was statistically significant t(117) = -3.41, p < .001, d = -0.64 with a medium effect size (d = -0.64; Cohen, 1988). The variances of the intercept were also significant for both women (ψ = 29.45, p < .001) and men (ψ = 21.67, p < .001), indicating that women and men showed meaningful individual variability around their respective group averages at baseline.

Table 3.11 Nested Comparison of Multiple Group Latent Growth Models for Financial Strain

Financial Strain	Parameters	χ^2	df	*p*-value	CFI	TLI	RMSEA	SRMR
Model 1: Means, Variances, Covariances, Residual Variances	12	63.31	42	0.018	0.912	0.937	0.092	0.162
Model 2: Means, Variances, Residual Variances	**11**	**63.37**	**43**	**0.023**	**0.916†**	**0.941**	**0.089**	**0.163**
Model 3: Means, Variances, Covariances	11	64.6	43	0.018	0.91	0.938	0.092	0.165
Model 4: Means and Variances	10	64.6	44	0.023	0.915	0.942	0.089	0.165
Model 5: Variances Only	8	161.64	46	< .001	0.521	0.687	0.206	1.439
Model 6: Means only	8	71.41	46	0.01	0.895	0.931	0.096	0.296

Note. n = 119. Model 2 (bolded) demonstrated the best fit and was retained for further analysis in aim 2. Missing data were estimated using full-information maximum likelihood (FIML). $\chi 2$ = chi-square goodness of fit test, CFI = Comparative Fit Index (Bentler, 1990), TLI = Tucker-Lewis Index (Tucker & Lewis, 1973), RMSEA = root mean square error of approximation (Steiger, 1990), and SRMR = standardized root mean-squared residual (Bentler, 1995). The best fitting model (bold) was chosen based on overall fit using the following cutoff recommendations: RMSEA < .06, p < .05; CFI/TLI ≥ .95; SRMR < .08. Once chosen, adjusted CFI and RMSEA values were determined using equivalence testing (Marcoulides & Yuan, 2017; Yuan et al., 2016) via the Dynamic Model Fit app, version 1.1.0 (McNeish & Wolf, 2020).

* Indicates null model should be rejected.

† Indicates excellent fit for the adjusted CFI value (≥ 0.915).

‡ Indicates excellent fit for the adjusted RMSEA value (≤ .063).

The mean of the slope factor for men's financial strain (μ = -0.61, $p < .001$) demonstrated a slight yet significant decrease, indicating that, on average, men's financial strain decreased over the six-month study period; no meaningful change was observed for women's financial strain. However, the slope variance for women's financial strain (ψ = 0.76, $p < .001$) was significant, reflecting meaningful individual variability in the rate of change in women's financial strain over time; no meaningful variability in change was observed for men's financial strain. The residual variances were significant for both genders, indicating that unexplained individual variability was observed across both groups (see Table 3.11). Finally, the relationship between the intercept and slope factors (i.e., covariance) was not significant (see Figure 3.5).

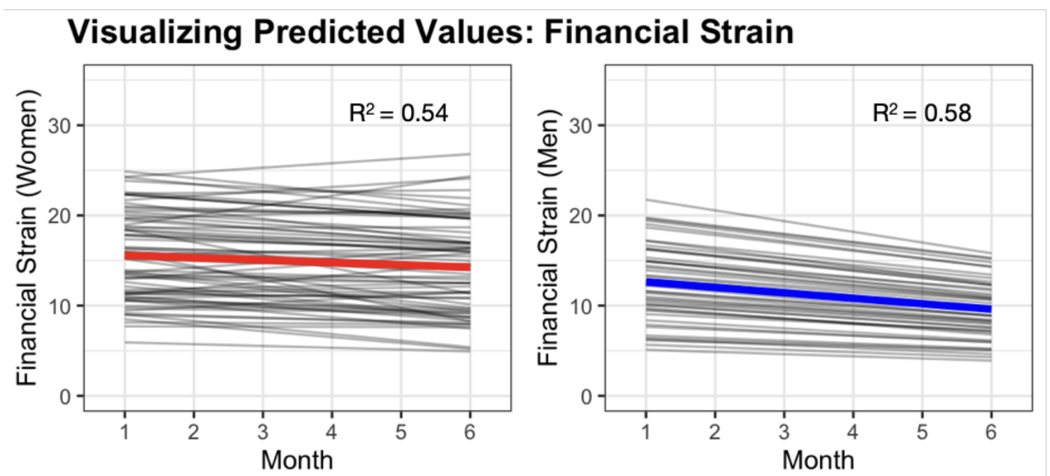

Figure 3.5 Predicted Financial Strain Trajectories for Women and Men over Six-Month Period

Note. R^2 was calculated as a measure of overall goodness of fit using the sum of squared residuals and the total sum of squares.

3.3.5 Social Support

Unlike the previous recovery capital variables, the best fit for the social support data was offered by the model where, in addition to freely estimating factor means, variances, and residual variances, covariances were also estimated separately across gender (χ^2 (43) = 63.37, p = .023; CFI = 0.916, RMSEA = 0.089, SRMR = 0.163; Model 1, Table 3.12). The means of the intercepts for social support were significant for both women (μ = 2.72, p < .001) and men (μ = 2.91, p < .001). Women's social support at baseline (i.e., intercept) was 0.18 points lower than men's, on average, however a Welch independent samples t-test (unequal variances) determined that this difference was not statistically significant. The variances of the intercept were also significant for both women (ψ = 0.77, p < .001) and men (ψ = 0.31, p = .002), indicating that both groups showed significant individual variability around their respective averages at baseline.

Table 3.12 Nested Comparison of Multiple Group Latent Growth Models for Social Support

Social Support	Parameters	χ^2	*df*	*p*-value	CFI	TLI	RMSEA	SRMR
Model 1: Means, Variances, Covariances, Residual Variances	**12**	**65.04**	**42**	**0.013**	**0.875**	**0.911**	**0.096**	**0.203**
Model 2: Means, Variances, Residual Variances	11	72.49	43	0.003	0.841	0.889	0.107	0.257
Model 3: Means, Variances, Covariances	11	73.9	43	0.002	0.833	0.883	0.109	0.215
Model 4: Means and Variances	10	85.29	44	< .001	0.777	0.848	0.125	0.299
Model 5: Variances Only	8	228.08	46	< .001	0.016	0.358	0.257	4.616
Model 6: Means only	8	90.18	46	< .001	0.761	0.844	0.127	0.329

Note. n = 120. Model 1 (bolded) demonstrated the best fit and was retained for further analysis in aim 2. Missing data were estimated using full-information maximum likelihood (FIML). $\chi 2$ = chi-square goodness of fit test, CFI = Comparative Fit Index (Bentler, 1990), TLI = Tucker-Lewis Index (Tucker & Lewis, 1973), RMSEA = root mean square error of approximation (Steiger, 1990), and SRMR = standardized root mean-squared residual (Bentler, 1995). The best fitting model (bold) was chosen based on overall fit using the following cutoff recommendations: RMSEA < .06, $p < .05$; CFI/TLI ≥ .95; SRMR < .08. Once chosen, adjusted CFI and RMSEA values were determined using equivalence testing (Marcoulides & Yuan, 2017; Yuan et al., 2016) via the Dynamic Model Fit app, version 1.1.0 (McNeish & Wolf, 2020).
† Indicates excellent fit for the adjusted CFI value (≥ 0.915).
‡ Indicates excellent fit for the adjusted RMSEA value (≤ .063).

Neither the women's nor the men's slope means were significant, indicating that neither group's social support changed meaningfully during the six-month period. However, the women's slope variance was significant ($\psi = 0.06$, $p = .003$), which reflected individual differences in the trajectory of women's social support over time. The slope means and variances for men were not significant, indicating that no meaningful change was observed in men's overall social support or in individual variability over time. The residual variances were significant for both genders, indicating that unexplained individual variability was observed across both groups (see Table 3.12). Finally, the social support model was the only one of the four where covariances were estimated separately and a significant, negative covariance estimate ($\sigma = -.14$, $p = .013$) was observed for the women's group. This negative covariance is paired with the women's positive slope (see Table 3.12), indicating that women who started with higher levels of social support experienced increases at a slower rate whereas women who started with lower levels experienced increases at a steeper rate (see Figure 3.6).

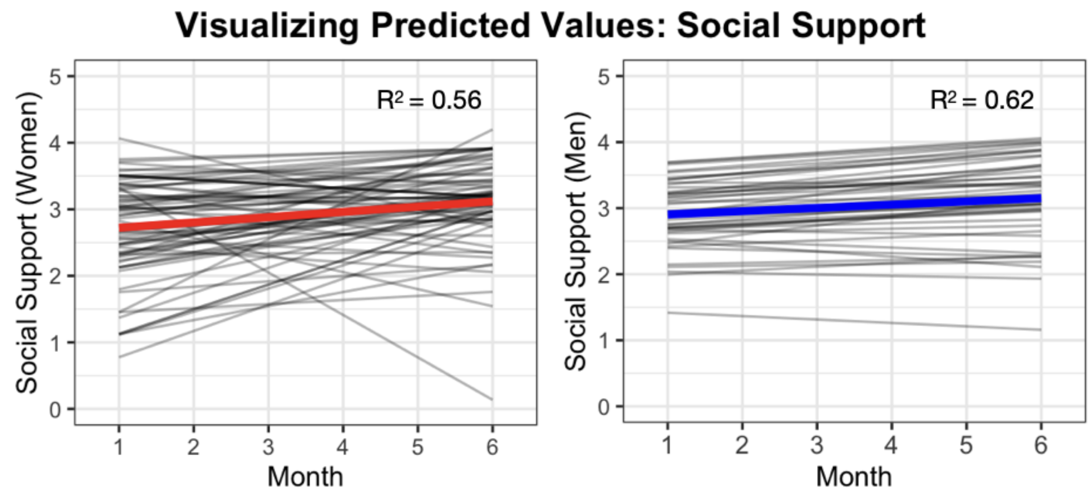

Figure 3.6 Predicted Social Support Trajectories for Women and Men over Six-Month Period

Note. R^2 was calculated as a measure of overall goodness of fit using the sum of squared residuals and the total sum of squares.

3.4 Random Forest Analyses

Three random forest feature-based classification models (full sample, women only, men only) were trained using the Scikit-learn library (Pedregosa et al., 2011) in Python (v. 3.10.2) to identify which features were most predictive of stable versus unstable recovery.

The features (predictors) were selected prior to the analysis based on domain expertise (see Table 2.4); these variables were cleaned and processed using Stata 15.1 and included sociodemographic characteristics and substance use/mental health covariates (previously described in Section 2.3.8). To incorporate the recovery capital variables in this model, parameter estimates from the multiple-group latent growth models (i.e., predicted values, residuals, and individual goodness of fit statistics) were included as features. The parameter estimates generated from models in Aim 1 were preprocessed using R, and predicted values as well as residuals were computed within the Latent Variable Modeling framework using "lavPredict" (Lavaan version 0.6-9; Rosseel, 2012) following guidelines provided by Hallgren et al. (2019). In a growth model, as opposed to a structural equation model (e.g., confirmatory factor analysis), "lavPredict" computes predicted (i.e., fitted) values for each individual based on the latent factors (i.e., latent intercept, latent slope) in the model and the observed values from the master data (Rosseel, 2012). The individual goodness of fit (GOF) statistics were calculated as an indicator of how well the multiple group latent growth curve model fit each individual as:

$$GOF_i = (y_{1i} - \hat{y}_{1i})^2 + (y_{2i} - \hat{y}_{2i})^2 + (y_{3i} - \hat{y}_{3i})^2 \ \ldots + (y_{6i} - \hat{y}_{6i})^2 \qquad (1)$$

where y_{ki} is the observed recovery capital variable at a given time point k for a given individual i, and $\hat{y}_{ki}$ is the predicted value derived from the multiple group latent growth model at a given time point k for a given individual i.

Each random forest model was trained and evaluated using a series of procedures as recommended by Stevens et al. (2020). The following data pre-processing procedures were applied to the data: after imputing missing data on the independent variables as described in the Methods section, multi-categorical (> 2 categories) nominal features were one-hot encoded. One-hot encoding is a process by which categorical variables are converted into dummy variables. Ordinal features (variables with ordered categories) were assigned integer values and labeled as integers. Other variables were treated as continuous. See Appendices C-D for an overview of the features included in each of the models. Next, the stratified k-fold procedure was used to split training and test data so that the label occurrence rate was consistent with that in the target set. The *k*-fold cross-validation procedure involves splitting the training dataset into *k* folds. The first *k*-1 folds are used to train a model, and the holdout *k*th fold is used as the test set. This process is repeated and each of the folds is given an opportunity to be used as the holdout test set. A total of *k* models are fit and evaluated, and the performance of the model is calculated as the mean of these runs. This, however, proved to be insufficient to ensure accurate results in the presence of significant imbalance and a small dataset due to the rarity of the "unstable" label. A randomized oversampling procedure was then utilized to address the skewed class proportions (i.e., imbalance) in the target variable; this procedure evens the occurrences of target 1's (i.e., unstable recovery) and target 0's (i.e., stable recovery) in the training data. These oversampled sets are used hereafter.

The design of a random forest model implies decisions over many hyper-parameters, such as the depth of the trees, number of trees, etc. (see section 2.4.3). A randomized grid search was therefore conducted to identify the best hyper-parameters for each model separately (see Table 3.13 for all tested parameter combinations). The model with the "best parameters" (i.e., the set of hyperparameters that maximized the accuracy over the test data in a 5-fold test run) was retrained on the training data and then used to predict the target using the test data.

Table 3.13 Hyperparameters Tested in Randomized Grid Search

Number of Trees	10	20	50	100	200	300	500	1000
Number of Features	30	46	92					
Number of Levels in Tree	2	3	4					
Bootstrapping	True	False						

Note. Bootstrapping refers to the sampling of observations either with (bootstrapping = True) or without (bootstrapping = False) replacement.

Finally, the accuracy score as well as the true/false positive/negative rates were measured and are presented below using a confusion matrix. Table 3.14 provides an overview of the performance metrics used to compare, assess, and determine the optimal model for each sample. Results for each of the three models are discussed in detail below.

Table 3.14 Random Forest Performance Metric, Formula, and Interpretation

Metric	Formula	Interpretation
Accuracy	$\frac{(TP + TN)}{(TP + TN + FN + FP)}$	Percentage of individuals correctly predicted to be in either unstable (TP) or stable recovery (TN) over the total number of individuals (TP+TN+FN+FP).
Precision/Positive Predictive Value	$\frac{(TP)}{(TP + FP)}$	Percentage of individuals correctly predicted to be in unstable recovery (TP) out of the total number of individuals predicted to be in unstable recovery (TP+FP).
Negative Predictive Value	$\frac{(TN)}{(TN + FN)}$	Percentage of individuals correctly predicted to be in stable recovery (TN) out of the total number of individuals predicted to be in stable recovery (TN+FN).
Sensitivity/ Recall	$\frac{(TP)}{(TP + FN)}$	Percentage of people correctly predicted to be in unstable recovery (TP) out of the total number of individuals who are actually in unstable recovery (TP+FN).
Specificity	$\frac{(TN)}{(TN + FP)}$	Percentage of people correctly predicted to be in stable recovery (TN) out of the total number of individuals who are actually in stable recovery (TN+FP).

Note. TP = true positive, TN = true negative, FP = false positive, FN = false negative.

In addition to generating predictions, the random forest performs an implicit feature selection analysis in which the model identifies a small subset of “strong variables” for its classification (Breiman, 2004, p. 2). This feature importance analysis can be interpreted as a selection of strong predictors of stable recovery. The outcome of the feature selection analysis is visualized as a relative ranking of features included in the model based on the importance score and will be discussed below for each model.

3.4.1 Full Sample Random Forest Model

Results from the randomized grid search indicated that the full sample model with the best hyperparameters was one with 500 estimators, a maximum depth of three, a maximum of 87 features, and bootstrapping (i.e., observations were sampled with replacement). Table 3.15 compares parameters across models.

Table 3.15 Comparison of Parameters across Random Forest Models

Parameter	Full Sample (n = 120)	Women (n = 73)	Men (n = 47)
Number of Trees	500	500	300
Number of Features in Data	87	82	82
Maximum Features Considered	87	9.05	41
Number of Levels in Tree	3	3	2
Bootstrapping	True	False	False

Each model achieved very high accuracy on the training data: 0.98 for the full sample, 1.0 for the women's model, and 1.0 for the men's model. However, each model showed decreasing accuracy in the (imbalanced) test data: 0.79 for the full sample, 0.73 for the women's model, and 0.8 for the men's model. This indicates some level of overfitting, which is not unexpected with a small and complex dataset such as this one.

The full sample classification model included 87 features and achieved a test classification accuracy of 0.79. Each feature's importance is ranked and visualized in Figure 3.7, and Figure 3.8 provides a visual representation of an individual decision tree sampled from the full sample random forest model. The three most important predictors of recovery status were *predicted depression at wave 4*, *length of stay in the sober living home* (with similar importance ~0.12, which can be interpreted as the "fractional importance" as the sum of the feature importances sums to 1 over the full feature set), and *predicted depression at wave 3* (with somewhat smaller importance at ~0.09). Beyond these three features, all remaining features showed a substantially smaller feature importance (< 0.06).

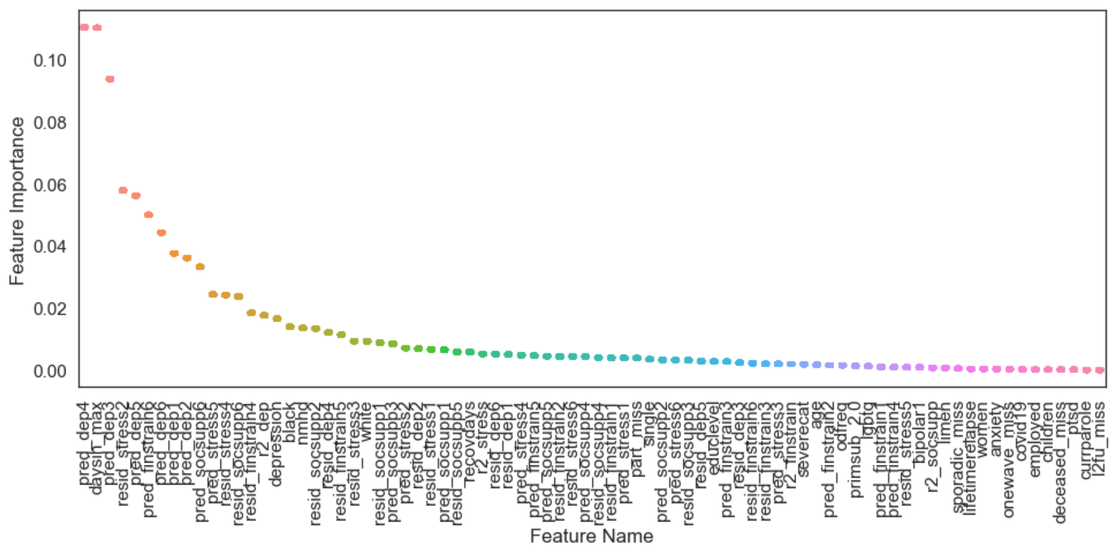

Figure 3.7 Feature Importances for Full Sample Random Forest Model

Note. Only feature importance values > 0 are visualized.

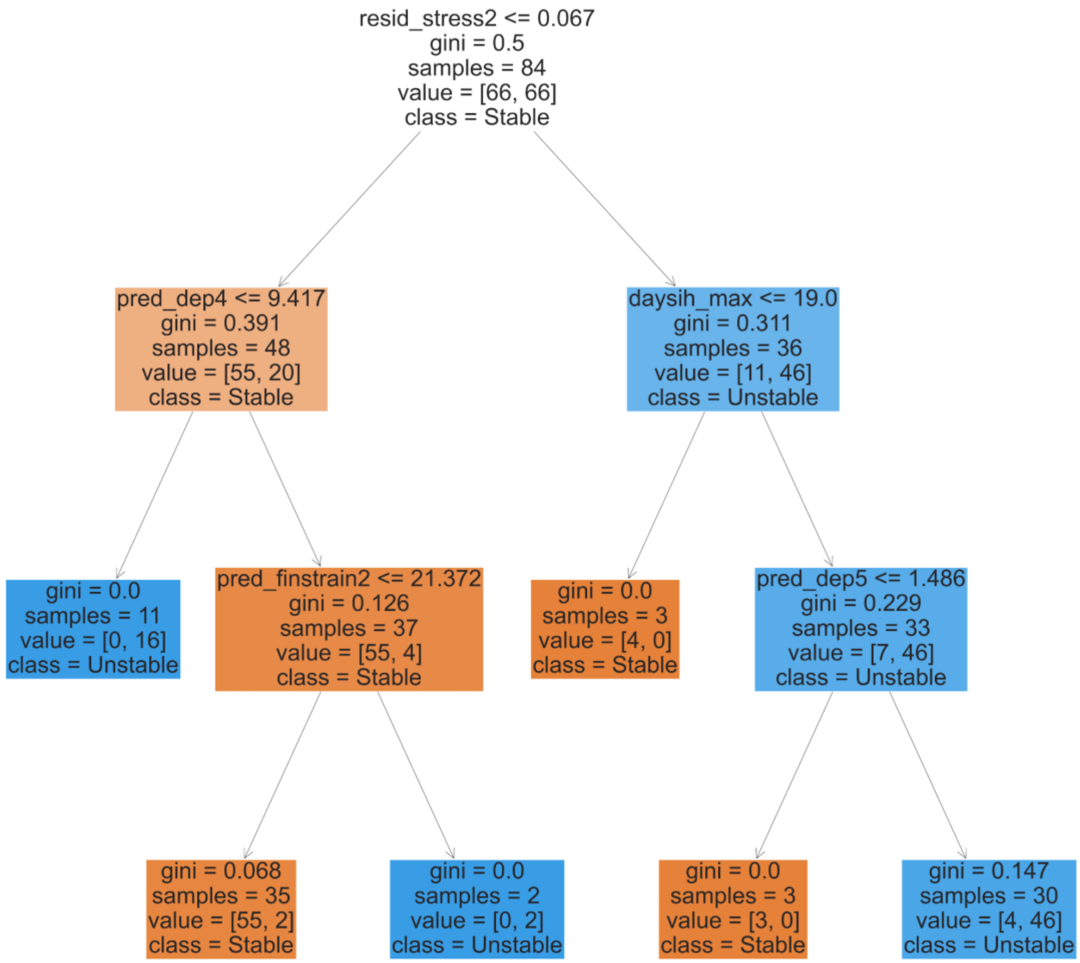

Figure 3.8 Individual Tree Sampled from Full Sample Random Forest

In this dendrogram (tree visualization; Figure 3.8), each node represents a point of split of the incoming (sub)sample. The name of the variable (i.e., residual stress at Month 2) used to split the sample and point of split are indicated at the top of each box, and the gini index represents the effectiveness of the split.

The effectiveness of the random forest classification model is visualized using a confusion matrix (see Figure 3.9), or a table that presents the summary of the prediction results. Out of the total number of people predicted to be in unstable recovery (target = 1; *n* = 18), the full sample model (79% accuracy) accurately classified 15 individuals (i.e., precision = 0.83). The model's negative predictive value was 0.67; meaning, the model accurately classified four out of the six individuals predicted to be in stable recovery. The model's average sensitivity (i.e., sensitivity = 0.88) was higher than the average specificity (i.e., specificity = 0.57), meaning the model performed better at correctly classifying individuals in unstable recovery compared to stable recovery (see Figure 3.9 for all model performance metrics). It is likely that the reason the model overpredicted the unstable class is due to the imbalance in the target data.

		Predicted		
		Stable	Unstable	
Actual	Stable	**4** (True Negative)	**3** (False Positive) X	**Specificity = 0.57** TN/TN+FP
	Unstable	**2** (False Negative) X	**15** (True Positive)	**Sensitivity/ Recall = 0.88** (TP/TP+FN)
		NPV = 0.67 (TN/TN+FN)	**Precision = 0.83** (TP/TP+FP)	**Accuracy = 0.79** (TP+TN) (TP+TN+FN+FP)

Figure 3.9 Confusion Matrix for Random Forest Model: Full Sample

Note. The summary of the classification model results is presented in each of the true positive (TP), true negative (TN), false positive (FP), and false negative (FN) cells. **Accuracy** is the percentage of individuals correctly predicted to be in either unstable (TP) or stable recovery (TN) over the total number of individuals (TP+TN+FN+FP). **Precision** is the percentage of individuals correctly predicted to be in unstable recovery (TP) out of the total number of individuals predicted to be in unstable recovery (TP+FP). **Negative Predictive Value (NPV)** is the percentage of individuals correctly predicted to be in stable recovery (TN) out of the total number of individuals predicted to be in stable recovery (TN+FN). **Sensitivity, or Recall**, is the percentage of people correctly predicted to be in unstable recovery (TP) out of the total number of individuals who are actually in unstable recovery (TP+FN). **Specificity** is the percentage of people correctly predicted to be in stable recovery (TN) out of the total number of individuals who are actually in stable recovery (TN+FP).

3.4.2 Gender-specific Random Forest Models

Each of the gender-specific models were filtered to only include women (n = 73) or men (n = 47). Due to the smaller sample sizes of the gender-specific samples, an analysis of the feature space was added to the model building procedures. The aim of the feature space analysis was to improve model performance by simplifying the feature space (i.e., decreasing the number of features). Results determined that certain features, which had been used in the full sample model, were not contributing meaningfully to the gender-specific models' performance (i.e., improving prediction accuracy while preserving low-bias and low-variance). For example, in the full sample model, all categories of race/ethnicity (i.e., white, non-Hispanic; Black, non-Hispanic; Hispanic; Native American, non-Hispanic; Other, non-Hispanic) were included as one-hot encoded features (dummy variables); however, in the gender-specific models, the feature importance analysis determined that only the white, non-Hispanic (versus all other) feature made a strong contribution to the classification. Therefore, white, non-Hispanic was the only race/ethnicity feature (dummy variable) retained in the final gender-specific models. See Appendices C-D for a comprehensive list of features included in each model.

3.4.2.1 Women Only Random Forest Model

Results from the randomized grid search indicated that the optimal women's model was one with 500 estimators, a maximum depth of three, a maximum of 9.05 (i.e., $\sqrt{n}$ features) features, and no bootstrapping (i.e., observations were sampled without replacement). This model achieved a test classification accuracy of 0.73 on the test set (see Figure 3.10 for confusion matrix). The three most important predictors of women's recovery status were *residual social support at wave 3*, *length of stay in* *sober living home*, and *residual depression at wave 1*. Residuals represent the difference between the observed recovery capital variable and the predicted value. In the current analysis, the residuals are an indicator of how well the predicted recovery capital data fit each individual at a given time point. Each feature's importance is ranked and visualized in Figure 3.10. And Figure 3.11 provides a visual representation of an individual decision tree sampled from the women's random forest model.

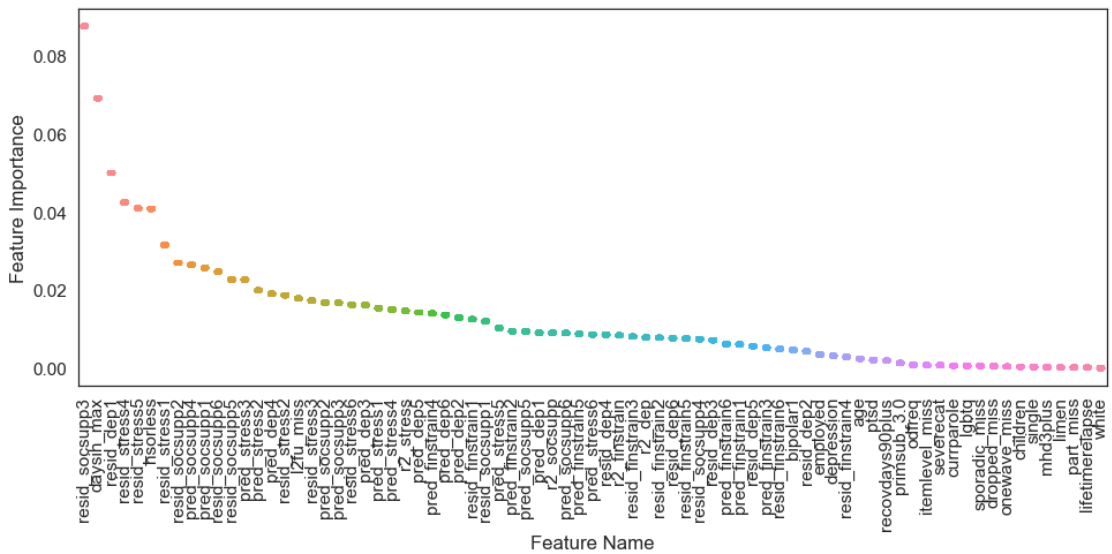

Figure 3.10 Feature Importances for Women's Random Forest Model

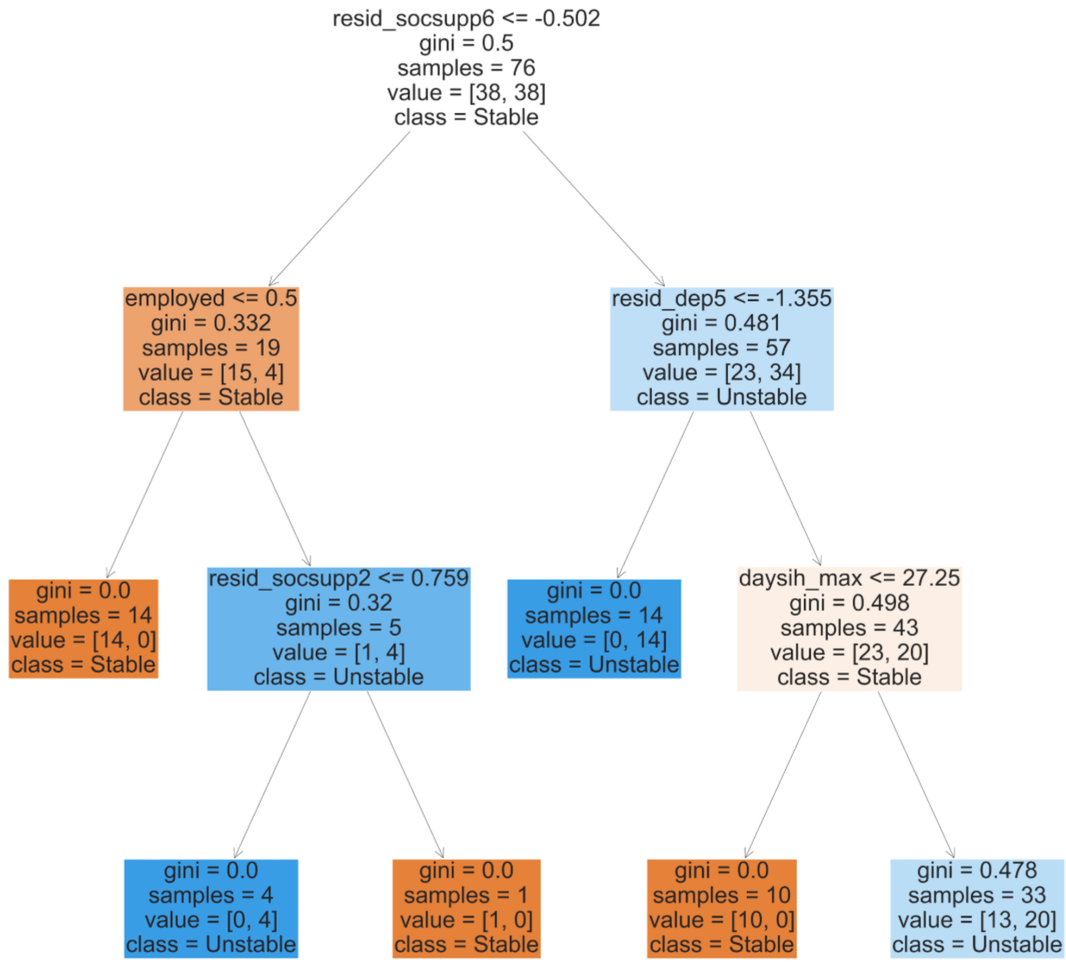

Figure 3.11 Individual Tree Sampled from Women's Random Forest

Out of the total number of people predicted to be in unstable recovery (target =1; n = 12), the model accurately classified nine individuals (i.e., precision = 0.75). The model's negative predictive value score was 0.67; meaning two out of the three individuals predicted to be in stable recovery were actually in stable recovery. The model's average sensitivity (i.e., sensitivity = 0.9) was higher than the average specificity (i.e., specificity = 0.4), meaning that the model was better at correctly classifying individuals in unstable recovery compared to stable recovery (see Figure 3.12 for all model performance metrics). Similar to the full sample model, the women's model likely overpredicted the unstable class due to the imbalance in the target data.

		Predicted		
		Stable	Unstable	
Actual	Stable	**2** (True Negative)	**3** (False Positive) X	**Specificity = 0.4** TN/TN+FP
	Unstable	**1** (False Negative) X	**9** (True Positive)	**Sensitivity/ Recall = 0.9** (TP/TP+FN)
		NPV = 0.67 (TN/TN+FN)	**Precision = 0.75** (TP/TP+FP)	**Accuracy = 0.73** (TP+TN) / (TP+TN+FN+FP)

Figure 3.12   Confusion Matrix for Random Forest Model: Women

Note. The summary of the classification model results is presented in each of the true positive (TP), true negative (TN), false positive (FP), and false negative (FN) cells. **Accuracy** is the percentage of individuals correctly predicted to be in either unstable (TP) or stable recovery (TN) over the total number of individuals (TP+TN+FN+FP). **Precision** is the percentage of individuals correctly predicted to be in unstable recovery (TP) out of the total number of individuals predicted to be in unstable recovery (TP+FP). **Negative Predictive Value (NPV)** is the percentage of individuals correctly predicted to be in stable recovery (TN) out of the total number of individuals predicted to be in stable recovery (TN+FN). **Sensitivity, or Recall**, is the percentage of people correctly predicted to be in unstable recovery (TP) out of the total number of individuals who are actually in unstable recovery (TP+FN). **Specificity** is the percentage of people correctly predicted to be in stable recovery (TN) out of the total number of individuals who are actually in stable recovery (TN+FP).

3.4.2.2 Men Only Random Forest Model

Results from the randomized grid search indicated that the men's model with the best parameters was one with 300 estimators, a maximum depth of two, a maximum of 41 features, and no bootstrapping (i.e., observations were sampled without replacement). Like the women's model, the men's model included 82 total features but achieved a test classification accuracy of 0.8. The three most important predictors of recovery status for men were residual *social support at wave 6*, *predicted social support at wave 4*, and *anxiety disorder diagnosis*. Although anxiety disorder diagnosis is included for symmetry with the reporting of the other models (i.e., three strongest predictors reported for each model), in this model the third feature (i.e., anxiety disorder diagnosis) does not have a significantly higher importance than the subsequent features, so it should not necessarily be considered "important." Each feature's importance is ranked and visualized in Figure 3.13, and Figure 3.14 provides a visualization of a decision tree sampled from the men's random forest.

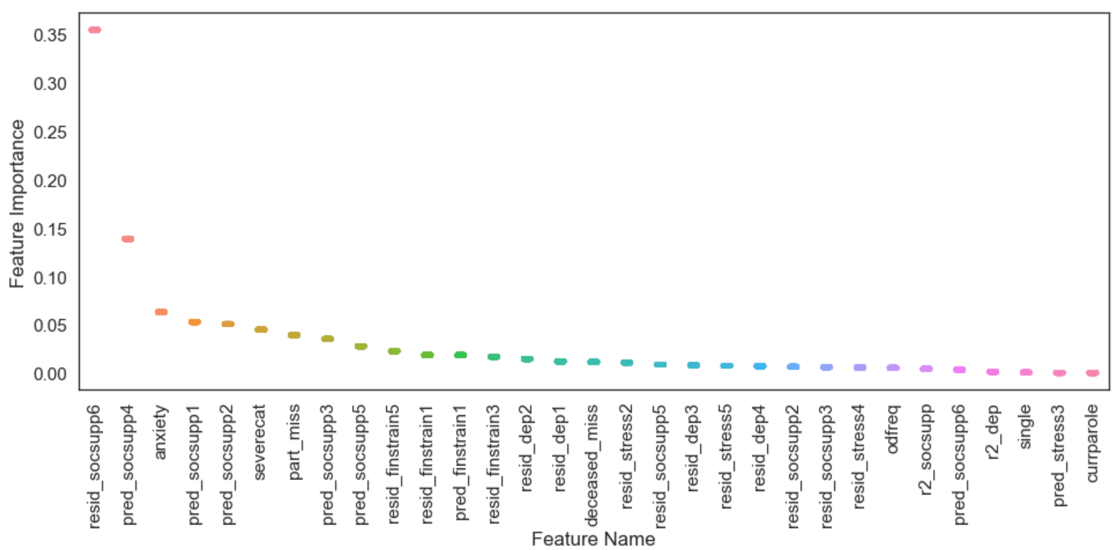

Figure 3.13 Feature Importances for Men's Random Forest Model

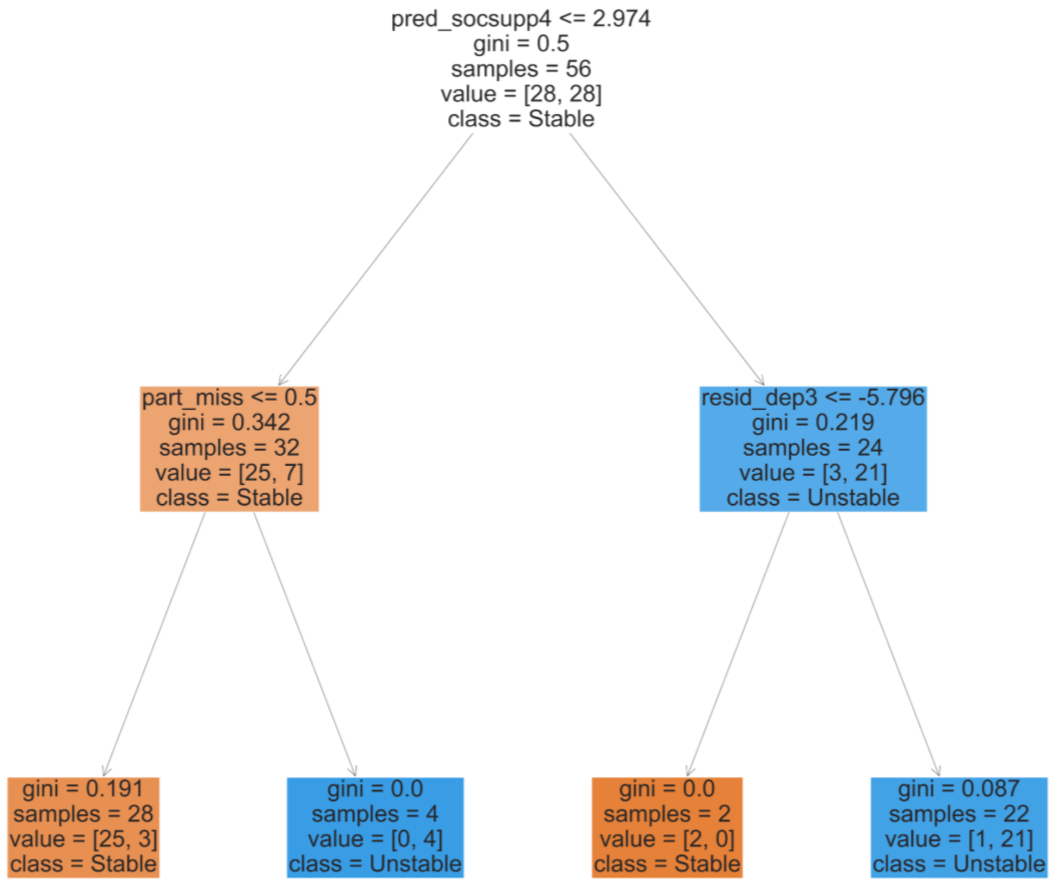

Figure 3.14 Individual Tree Sampled from Men's Random Forest

Out of the total number of people predicted to be in unstable recovery (target =1; n = 9), the model accurately classified seven individuals (i.e., precision = 0.78). The model's negative predictive value score was 1.0; meaning the one individual who was predicted to be in stable recovery was actually in stable recovery (see Figure 3.15 for all model performance metrics). The men's model demonstrated the most instability of the three models; meaning, with each change to the hyperparameters, the men's model exhibited more variability compared to either the full sample model or the women's model. This was likely due to the smaller amount of data available in the men's dataset (*n* = 47). While the performance varied with adjustments to the hyperparameters, the feature importance analysis remained stable; meaning, the strongest predictors remained constant despite changes in the model's performance. The stability in the feature importance analysis implies that the variability in model performance was most likely stochastic (random) in nature rather than due to any sort of meaningful bias.

		Predicted		
		Stable	Unstable	
Actual	Stable	**1** (True Negative)	**2** (False Positive) **X**	**Specificity = 0.33** TN/TN+FP
	Unstable	**0** (False Negative) **X**	**7** (True Positive)	**Sensitivity/ Recall = 1.0** (TP/TP+FN)
		NPV = 1.0 (TN/TN+FN)	**Precision = 0.78** (TP/TP+FP)	**Accuracy = 0.8** (TP+TN) (TP+TN+FN+FP)

Figure 3.15 Confusion Matrix for Random Forest Model: Men

Note. The summary of the classification model results is presented in each of the true positive (TP), true negative (TN), false positive (FP), and false negative (FN) cells. **Accuracy** is the percentage of individuals correctly predicted to be in either unstable (TP) or stable recovery (TN) over the total number of individuals (TP+TN+FN+FP). **Precision** is the percentage of individuals correctly predicted to be in unstable recovery (TP) out of the total number of individuals predicted to be in unstable recovery (TP+FP). **Negative Predictive Value** (NPV) is the percentage of individuals correctly predicted to be in stable recovery (TN) out of the total number of individuals predicted to be in stable recovery (TN+FN). **Sensitivity**, or Recall, is the percentage of people correctly predicted to be in unstable recovery (TP) out of the total number of individuals who are actually in unstable recovery (TP+FN). **Specificity** is the percentage of people correctly predicted to be in stable recovery (TN) out of the total number of individuals who are actually in stable recovery (TN+FP).

Table 3.16 reports the sensitivity, specificity, and accuracy metrics for all three models. The full sample random forest and the men's random forest predicted recovery status with similar levels of accuracy (full sample, 0.79; men, 0.8); however, the men's random forest classifier's specificity fell 0.24 points below that of the full sample model. All models demonstrated high levels of sensitivity, meaning they all were able to accurately and consistently predict unstable recovery. All of the models had higher levels of sensitivity compared to specificity, meaning that they are more likely to overpredict unstable recovery (i.e., some of the individuals who are predicted to be in unstable recovery are actually in stable recovery).

Table 3.16 Comparison of Sensitivity, Specificity, and Accuracy Metrics on Test Data across Random Forest Models

Model	Sensitivity	Specificity	Accuracy
Full Sample	0.88	0.57	0.79
Women	0.9	0.4	0.73
Men	1.0	0.33	0.8

Note. **Sensitivity** is the percentage of people correctly predicted to be in unstable recovery (TP) out of the total number of individuals who are actually in unstable recovery (TP+FN). **Specificity** is the percentage of people correctly predicted to be in stable recovery (TN) out of the total number of individuals who are actually in stable recovery (TN+FP). **Accuracy** is the percentage of individuals correctly predicted to be in either unstable (TP) or stable recovery (TN) over the total number of individuals (TP+TN+FN+FP).

Chapter 4

DISCUSSION

Women and men develop and recover from SUD differently; however, little is known about gender-specific attributes that differentiate women's and men's recovery beyond the initiation stage of abstinence (e.g., treatment) or in recovery housing contexts. This study contributes novel evidence about gender differences in recovery capital for individuals residing in recovery housing to the growing body of literature on gender differences in SUD recovery. Findings showed that, at baseline, women presented with more trauma and co-occurring mental health disorders, made less money, and reported greater financial strain, stress, and depressive symptomatology compared to men. Results from the multiple-group latent growth curve models suggested that specific forms of women's and men's recovery capital, namely perceived stress and financial strain, differed in terms of rate of change over a six-month period. Results from the random forest classifier models suggest that, for both women and men, recovery capital indicators mattered more in predicting recovery status, compared to other substance use, mental health, and sociodemographic variables. Social support, one type of recovery capital, was the strongest predictor of both women's and men's recovery status; length of stay in recovery residence was a strong predictor for women but not for men; and anxiety disorder diagnosis was a strong predictor for men, but not for women.

4.1 Summary of Findings

4.1.1 Exploratory Analyses

Compared to men, women in this study presented with more trauma (i.e., PTSD diagnoses) and co-occurring mental health disorders, made less money, and reported greater financial strain, stress, and depressive symptomatology at baseline. Previous research on gender differences in SUD treatment show that women are more likely than men to present with complex SUD diagnoses (McHugh et al., 2018) and trauma (Hecksher & Hesse, 2009; Lotzin et al., 2019) and experience more severe impairment related to psychiatric functioning (Foster et al., 2016; Hernandez-Avila et al., 2004; McHugh et al., 2013; Wu et al., 2010). Current findings extend previous research by demonstrating that the challenges women face during treatment persist beyond the treatment stage into early recovery and in recovery housing contexts.

This study's findings also support previous research from sober living homes studies using women-only samples. Women residing in sober living homes often experience mental health problems, financial strain, and housing insecurity (Edwards et al., 2017; Krentzman et al., 2022) and may derive particular benefit from trauma-informed recovery environments (Edwards et al., 2021). The current study's findings related to depressive symptomatology, perceived stress, and financial strain further amplify the call made by previous researchers (Cano et al., 2017; Edwards et al., 2017; Krentzman et al., 2022) to integrate services in recovery housing that are responsive to the specific needs of women in recovery (i.e., gender-responsive). Such programming may include trauma-informed services (Cano et al., 2017; Edwards et al., 2017) as well as policies designed to address the financial and mental health needs of women in recovery (Krentzman et al., 2022).

4.1.2 Aim 1: Multiple-Group Latent Growth Curve Models

Multiple-group latent growth models characterized women's and men's recovery capital trajectories while residing in and transitioning from sober living homes over a six-month period during early recovery. Results from exploratory and preliminary analyses indicated that women and men do not differ significantly in terms of internalized stigma or mindfulness. In terms of depression and social support, neither women's nor men's trajectories differed significantly at baseline or over time. Women's social support did, however, demonstrate a significant negative covariance, which suggested that women who started with higher levels of social support at baseline experienced increases at a slower rate compared to women who started with lower levels and experienced increases at a steeper rate (see Figure 3.6). The primary differences in women's and men's recovery capital trajectories were observed in terms of perceived stress and financial strain.

4.1.2.1 Perceived Stress

While men's initial levels of perceived stress ($M = 6.54$) started lower than women's ($M = 7.22$), women experienced a significant decrease in stress over time to the point where both women and men demonstrated relatively similar levels at month six (see Figure 3.4). The significant negative slope in the women's model suggests that stress decreases with additional time in recovery for women residing in and transitioning from sober living homes. Research on gender differences in SUD recovery suggests that one of the drivers of women's substance use is stress. Women are more likely to use drugs and alcohol to cope with stress, negative affect, and/or trauma during active use (Haseltine, 2000); however, an improvement in women's social context post-treatment has been shown to be associated with a decrease in stress

(Timko et al., 2005). One of the goals of a sober living home is to provide individuals in recovery with a supportive sober community (i.e., supportive social context). Given previous associations identified between improvements in women's social contexts and decreases in stress, current findings suggest that recovery residences may positively impact women's ability to cope with stress in recovery by way of their social context. Future research could explore this association by first replicating the current analysis with a larger sample and then extending it using a mediation model where time in sober living home predicts change in social support, which in turn predicts change in stress, controlling for factors such as time in recovery, mental health disorder diagnoses, and other key substance use and sociodemographic variables.

4.1.2.2 Financial Strain

While men's financial strain decreased with time, no change was observed in terms of women's financial strain, which started higher and stayed higher than men's throughout the duration of the study. Thus, this finding implies that women residing in and transitioning from sober living homes experience greater and more persistent financial strain than men. Financial strain can be conceptualized as a form of chronic economic stress that, with time, contributes to allostatic load,[2] compromising one's resistance to illness and disease with age (Hayward et al., 2000; Kahn & Pearlin, 2006). Because women are overrepresented among those living in poverty and, despite similar levels of education, are paid less than men in the public sector while shouldering more responsibilities in the domestic sphere, it is likely that financial

---

[2] Allostatic load refers to the cumulative burden of chronic stress and life events (Guidi et al., 2021).

strain manifests in women in ways that are distinct from men (Shippee et al., 2012; U.S. Bureau of Labor Statistics, 2021). In their longitudinal study examining financial strain in women during middle and older adulthood, Shippee et al. (2012) found financial strain to be associated with rapid declines in women's health, especially for those who reported recurrent strain. In the current study, women experienced greater financial strain than men that persisted at relatively the same rate across time. Given that women are more likely than men to use drugs and alcohol to cope with stress (McHugh et al., 2018), financial strain may represent a gender-specific pathway to relapse for women in recovery. Together, these findings suggest that gender-responsive efforts should be taken to support women in recovery as they may be particularly vulnerable to the negative and cumulative effects of financial strain.

4.1.3 Aim 2: Random Forest

Three random forest classification models were built to predict stable versus unstable recovery for individuals (n = 120; 61% women) residing in and transitioning from sober living homes in Delaware. The purpose of the full sample model was to explore the relative importance of gender compared to other key features. The goal of the gender-specific models was to explore, compare, and contrast the strongest predictors of women's and men's recovery status. Results from the full sample model showed that gender was among the least predictive of recovery status relative to other features included in the model. However, it is important to note that many of the features included in the full sample model, especially those relating to recovery capital, were demonstrated to be related to gender in Aim 1 (e.g., financial strain, perceived stress) and proved to be strong predictors in the full sample Random Forest.

Therefore, it is possible that the importance of gender in the full sample model may be masked by the importance of other related features.

The overall accuracy for all three models was > 70%. While the sensitivity and specificity scores varied across models, sensitivity was consistently higher compared to specificity; meaning, the models all tended to overpredict the unstable recovery class. A similar trend was observed in a study using tree methods (logistic regression, decision tree, random forest) to predict opioid use disorder in a much larger sample (*n* = 42,324) from the National Survey on Drug Use and Health (NSDUH); the authors found that sensitivity was consistently higher than specificity across models (Wadekar, 2020). In both the current study and Wadekar's (2020) study, it is preferable for models to have a higher sensitivity score because accurately identifying people at risk and intervening with appropriate supports can translate to saved lives in the field of substance use. The overall accuracy of the prediction despite the study's small sample size reinforces the promise gender-specific machine learning methods show in recovery contexts. Employing such methods to predict recovery pathways in community-based settings can help practitioners tailor SUD treatment programming to meet the differential needs of both women and men.

4.1.3.1 Women

4.1.3.1.1 Social Support

The top three predictors of women's recovery status included social support, length of stay in sober living home, and depression. The current analysis included items that characterized emotional and informational social support; typically, social support is defined as a variety of helping behaviors that fall into one of three categories: emotional (e.g., encouragement), tangible (e.g., driving a friend to an appointment), and informational (e.g., advice, guidance; Barrera & Ainlay, 1983). Previous research suggests that emotional and informational support are associated with women's positive recovery outcomes such as treatment completion (Tracy et al., 2010), self-care (Brooks et al., 2017) and improved coping skills related to recovery (Tracy et al., 2010). Moreover, sober living homes have been shown to serve as a social context that fosters stability, safety, and sobriety (i.e., "safe zones;" Brooks et al., 2017, p. 79; Polcin, 2008) while promoting social support specific to recovery (Polcin, 2001). While the current analysis cannot determine definitively whether sober living homes were the primary mechanism of women's social support, previous research suggests it is possible. Future research could explore this relationship by examining social support over time as well as assessing women's social networks to better understand the potential roles sober living homes play in promoting social support, the associations between social support and other key recovery capital variables, as well as the communities from which women derive social support (see Brereton et al., 2014; Stevens et al., 2015; Tracy et al., 2010 for examples of social network analysis frameworks in sober living homes).

4.1.3.1.2 Length of Stay

Length of stay was the second strongest predictor of women's recovery status. Previous studies on sober living homes have established that length of stay matters in predicting improved outcomes, however, findings are mixed in terms of the amount of time that provides the optimum treatment effect. While the National Institute on Drug Abuse (2020) recommends a minimum stay of 90 days (three months), studies from large samples of Oxford House residents suggest a minimum stay of six months for maximum benefit (Jason et al., 2007, 2016). Mahoney et al. (2021) found that perceptions of social environments (i.e., support, involvement, practical orientation, order and organization) were significantly related to length of stay among residents, and other studies found that length of stay was associated with improved outcomes such as sobriety and employment (Jason et al., 2007). Inversely, earlier exit from sober living homes have been shown to predict more severe SUD and psychiatric problems (Harvey et al., 2016); reasons residents leave were often associated with financial stability or a desire for independence (Chavira and Jason, 2021). Given the current study's findings regarding women's financial strain and the relationship identified between financial instability and leaving sober living homes early (Chavira & Jason, 2021), future research could further explore financial strain as a potential moderator of women's length of stay.

While the studies discussed above lay an important foundation, their findings are based on samples wherein men make up the majority of participants (Aase et al., 2014; Chavira & Jason, 2021; Harvey et al., 2016; Jason et al., 2007); therefore, it is important to look to gender-specific evidence to confirm whether factors identified for men similarly predict length of stay for women. In gender-specific studies, age was a consistent predictor of longer stays in home for both women and men (Bishop et al., 1998; Edwards et al., 2022). Krentzman et al. (2022) further explored the relation between reasons why women leave sober living homes and age; the authors found that women in emerging adulthood (i.e., age 18-29) were more likely to break the house rules and be asked to leave than older women. In the same study, they found that the number of mental health disorders women reported was highly predictive of relapse and rule-breaking (Krentzman et al., 2022). The current study's findings support the association between women's recovery status and length of stay identified by previous literature and extends this line of research by including longitudinal (i.e., time-varying) recovery capital data in its prediction of relapse and/or early exit from sober living homes (i.e., unstable recovery). Future research predicting women's exit as well as their retention should examine recovery capital alongside key sociodemographic and mental health variables (e.g., age, number of mental health disorders).

4.1.3.1.3 Depression

Depression symptoms were the third strongest predictor of recovery status for women, but not for men. This finding is supported by previous research, which suggests that depression plays a stronger role in recovery among women relative to men. In a study examining gender-specific effects of comorbid depression and anxiety on the propensity to drink in negative emotional states, Karpyak et al. (2016) found that women with alcohol use disorder, compared to men with alcohol use disorder, demonstrated higher rates of lifetime major depression, substance-induced depression, anxiety disorder, and post-traumatic stress disorder (PTSD) and were more likely to drink alcohol when experiencing negative emotion. These findings reinforce previous evidence about motivators for women's substance use, which shows that women are more likely to use drugs or alcohol as a coping resource in response to negative affect or stress (Haseltine, 2000; Liu & Kaplan, 1996). In terms of the relationship between depression and women's recovery over time, Oliva et al. (2018) found that depression persisted over time and predicted a higher likelihood of women's relapse at a 12-month follow-up. Given women's increased likelihood to use drugs and alcohol to cope with negative emotions such as stress and depression, it may be important for women in sober living homes to be connected to mental health services to help alleviate symptoms related to mood disorders (e.g., depression, bipolar disorder).

4.1.3.2 Men

4.1.3.2.1 Social Support

Social support and anxiety disorder diagnosis were the strongest predictors of men's stable versus unstable recovery. The majority of research on the role of social support in predicting relapse has been conducted with mostly male samples (Aase et al., 2014; Chavira & Jason, 2021; Harvey et al., 2016; Jason et al., 2007) and shows that it is a key predictor of improved outcomes (i.e., sobriety, employment) during recovery. The current study suggests that social support may play similarly important roles in predicting both men's and women's recovery. While gender differences research is limited in this area, a recent study by Smith et al. (2018) explored the extent to which gender moderates the relationship between social support and risk of relapse in a collegiate recovery community (*n* = 148; 48% women; *M* age = 27) and found that men who reported higher rates of social support were found to experience an increased risk of relapse, whereas women who reported higher levels of social support exhibited a lower risk of relapse. Results from the current study's random forest models provide information about the strength of predictors, however, they do not provide information about the directionality of the relationships. Moreover, Smith et al.'s (2018) study suggests that the association between relapse and social support functions differently for women and men, however, more gender-specific research is needed to validate these findings and further explore both the direction and the strength of such relationships for women and men.

Previous research examining the role of social support in sober living homes has found that residents report receiving support from multiple sources. In their study examining the role of social support and spirituality in sobriety among a sample of men residing in Oxford Houses, Nealon-Woods et al. (1995) found that men received social support specific to their sobriety both within Oxford homes as well as from their broader 12-step communities. Because the current study did not include a social network component, it is not possible to determine whether men derived social support from their sober living home community or from other sources. Moreover, length of stay in home, a potential indicator of the sober living home "treatment effect," did not emerge as a strong predictor of men's recovery status as it did for the women. In order to understand the interactive role of social support and sober living homes in preventing relapse, it is critical that future research conduct gender-specific analyses on data that characterizes both social support and residents' social networks. Future research in this area might follow the example of Stevens et al.'s (2015) study that investigated social support and network relationships for residents of sober living homes using several indicators of social support and social networks.

In the same way women's social support may function differently from men's, women's and men's within-group needs related to social support may also differ. For example, in addition to the general challenges faced by men in recovery, sexual minority men navigate risk factors that are unique to sexual minorities (e.g., stigma, prejudice, discrimination; Mericle et al., 2019). Such risk factors can contribute to hostile social environments, increasing the risk for negative outcomes during recovery (e.g., relapse; Meyer, 2003). Support and solidarity from others within the gay community can serve as a critical buffer against sexual orientation-related stress

(Kelly et al., 2014; Meyer, 2003; Toomey et al., 2018). In their article examining recovery outcomes among sexual minority men, Mericle et al. (2018) recommended recovery housing that is specifically designated for sexual minority populations so that operators can focus resources to address group-specific needs. This is, essentially, the LGBT male equivalent of gender-responsive treatment contexts for women - a safe and stable place for recovery that is responsive to the group-specific needs of sexual minority men.

4.1.3.2.2 Anxiety Disorder Diagnosis

Anxiety disorder diagnosis was the other primary predictor of men's recovery status in the random forest model. While more women than men present with co-occurring depression and anxiety diagnoses at SUD treatment (Oliva et al., 2018), we are beginning to see differences in the way such symptomatology manifests between genders. In two majority men samples, Harvey et al. (2016) found that more severe SUD and psychiatric problems were related to earlier exits over a period of one year; whereas lower anxiety was related to longer stays in sober living homes (Aase et al., 2014).

Men who struggle with anxiety symptoms during recovery may need specific support to develop adaptive methods of coping. Previous research has found that experiential avoidance (i.e., propensity to avoid, escape, or change one's experience of trauma; Hayes et al., 1996) acts as a mechanism linking depression and anxiety with harmful health behaviors among men in SUD treatment. One study examining the compulsive sexual behavior of men in SUD treatment found depression and anxiety symptomatology to be indirectly associated with compulsive sexual behaviors through experiential avoidance. The authors recommended that recovery treatment settings help men in SUD recovery develop more adaptive methods of processing trauma, as opposed to escaping them (Brem et al., 2017). Future research should additionally assess anxiety symptoms as an indicator of recovery capital over time alongside stress, depression, financial strain, and social support.

4.2 Strengths & Limitations

Findings should be interpreted both in light of the current study's strengths and its limitations. One key strength of the study was its use of a community-based longitudinal dataset. Community partners initially requested that the parent study be conducted in an effort to better understand the outcomes of individuals residing in affiliated sober living homes; findings from this study therefore have potential to inform community-based recovery environments. Within the parent study, recovery capital data were collected over time. Data captured across time helps to inform the field's understanding of the rate and shape of change in recovery capital; this matters because recovery is a developmental process and the majority of research on gender differences in SUD and recovery focuses on the initiation stage of recovery (i.e., treatment, initiation of abstinence behaviors). This work contributes to the burgeoning area of research on sober living homes, which extends beyond the treatment stage or initiation stage of abstinence. Research that follows people in recovery across time can identify intervention strategies that are best tailored to different stages of recovery (e.g., early versus sustained, long-term recovery).

A methodological strength of the study was the gender-specific approach that was used to analyze the data. Throughout every stage of the study from the exploratory analyses to the primary aims, women's and men's outcomes were compared, contrasted, and analyzed separately. This approach was further strengthened by the use of a traditional multiple-group growth curve modeling approach coupled with a novel machine learning method to explore the gender-specific relationships between recovery capital and recovery status. The latent growth modeling technique allowed for the inclusion of gender-specific recovery capital growth parameters to be included as features in the random forest prediction. While primarily a strength of the study, the random forest model is relatively new to the substance use literature and, consequently, there are few studies with which to compare results. Nonetheless, machine learning models, like random forest, hold great promise in the recovery field, as well-performing models can be used to predict who is most vulnerable as well as who is most likely to benefit from treatment. Additionally, feature importance analyses are capable of identifying variables that can help inform and tailor treatment programming.

The dataset was not without its limitations. For example, the sample size was small (n = 120) compared to most machine learning studies (see Wadekar, 2020; n = 42,324), yet comparable to other community-based samples in the sober living home literature, which range from n = 10 (Chavira & Jason, 2021) to n = 897 (Aase et al., 2014; Jason et al., 2007) with most studies falling closer to n = 200 (Edwards et al., 2022; Harvey et al., 2016; Hunter et al., 2012; Mahoney et al., 2021; Polcin et al., 2010). The issue of sample size was further complicated by the gender-specific method utilized in Aim 2, which stratified the sample by gender, further decreasing the sample size by 39% for the women's model and 61% for the men's model. Nonetheless, random forest has been shown to work well with small datasets and therefore proved an appropriate analytic choice despite the study's sample size-related challenges. Future research should replicate this analytic approach (i.e., gender-specific random forests) with larger and more diverse samples in terms of race, culture (e.g., ethnicity, acculturation, nativity), social identity (e.g., LGBTQ status), and socioeconomic status to determine whether gender-specific trends remain stable.

The imbalance (i.e., skewness) in the target variable (i.e., recovery status) was an added challenge for the gender-specific analyses in Aim 2. To address the imbalance in the target data, a randomized oversampling procedure was utilized where random samples from the minority class (i.e., unstable recovery) were selected with replacement. While this procedure addresses imbalances in the data, it also introduces the risk of overfitting as there is a possibility that a single case may be selected multiple times.

Despite prevention efforts, the study suffered from higher-than-recommended levels of attrition; by month six (i.e., 52%), the retention rate fell below the recommended rate for the field (i.e., 60%; Hansten et al., 2000). Although multiple steps were taken to prevent and address the threat of attrition (e.g., missing data analysis, attrition analyses, statistical imputation), the risk remains that certain characteristics of participants who dropped out at later waves differed from those who stayed in the study (e.g., White, non-Hispanic vs. Other, LGBTQ). Such systematic differences potentially introduce bias into the results, though not necessarily. The greatest amount of dropout occurred amongst people who moved out of the sober living homes. Beyond improving retention efforts during the data collection phase, two analytic steps are recommended to improve the validity of findings in longitudinal sober living homes studies. First, if a participant moves out, data should be collected on their reason for leaving (i.e., conditions of exit). Reason for leaving could then be used in a survival analysis to predict dropout as well as a covariate in the longitudinal analyses. Second, future research should create a plan for documenting missing data patterns as they occur as part of the data collection procedures to better characterize the nature of potential bias. Missing data categories identified in the current study can be used to inform future sober living homes studies in terms of both prevention strategies as well as missing data analyses.

Finally, while a conservative parameter of early exit (i.e., < 30 days) was chosen as a partial indicator of recovery status, there is no definitive way to be certain that those who left the home after staying less than 30 days indeed experienced instability in their recovery. This indicator was chosen based on the literature to characterize early exit because it fell five months below Jason et al.'s (2007, 2016) recommendation and two months below than the National Institute on Drug Abuse's (2020) recommendation. Research also shows that individuals with shorter stays are more likely to leave due to rule-breaking or relapse (Krentzman et al., 2022). In order to improve on this measure, more research is needed regarding optimum length of stays in sober living homes and whether or how such durations differ by gender or other factors. The average stay in the current study for women was five months whereas, for men, it was closer to nine months. Future gender-specific studies could build on this measure by examining minimum lengths of stay for women versus men. Furthermore, in the current study, length of stay was a strong predictor of women's recovery status, but not men's. These distinguishing factors matter because women are more likely to face challenges in their recovery related to childcare responsibilities, financial strain, and domestic violence and, therefore, may require length of stay recommendations that are responsive to their specific needs. Such gender-responsive recommendations should be tailored using a health equity lens in order to account for the varying degrees of vulnerability experienced by certain groups (e.g., sexual minority men, women with children in their custody, women with co-occurring domestic violence, homelessness).

4.3 Conclusions and Implications

The current study contributes gender-specific evidence regarding recovery outcomes for women and men residing in sober living homes to a growing literature base on gender differences in SUD recovery. While social support weighed equally in importance for both women's and men's recovery, it is important to contextualize this finding in light of other factors. Findings reinforce what is already known about women in recovery; that is, the needs, experiences, and strengths of women differ in ways that are distinct from men.

4.3.1 Addressing Recovery Capital in Recovery Housing

Social support, as a key source of recovery capital, was the strongest predictor of recovery status for both women and men. At face value, these findings suggest that social support impacts women's and men's recovery in similar ways. However, when the many other differences in women's and men's recovery pathways are taken into account, the interpretation shifts. Women in this study, like in many before it, presented with more trauma and co-occurring mental health disorders; made less money; and, reported greater financial strain, stress, and depressive symptomatology compared to men across a six-month period. Taken together, these findings suggest that women experience greater barriers to recovery in terms of economic strain and psychological functioning during early stages of recovery yet are still benefiting from social support in similar ways to men. Contextualizing women's social support within the broader constellation of gender-specific barriers women in recovery face is critical in order to build a more nuanced understanding of the role social support plays among women in recovery.

In addition, women experienced worse financial strain than men that did not lessen with time. This is likely due to both the larger societal inequities women face in terms of economic strain (e.g., gender wage gap) intersecting with the particular challenges women face in recovery (e.g., co-occurring mental health disorders, experiences with trauma/domestic violence, caregiving responsibilities/custody issues). Efforts should be made to address both the underlying economic disparities that contribute to financial strain as well as the ways in which economic stress manifests physically and psychologically for women in recovery. Examples of such efforts could include an integrative system of care (e.g., gender-responsive programming) that addresses economic strain by providing childcare for individuals who are primary caregivers, stress reduction techniques (e.g., mindfulness-based stress reduction) to cope with psycho- and physiological effects of financial strain, and educational resources (e.g., financial literacy) to improve knowledge about finances (e.g., budgeting). It is important to note that as part of a gender-responsive framework, these offerings should be culturally responsive (Covington & Bloom, 2017).

Finally, depressive symptomatology was among the strongest predictors of women's recovery status whereas anxiety disorder diagnosis was among the strongest predictors for men. These findings are consistent with previous literature, which show that mental health conditions affect both women and men, but anxiety is more strongly associated with men's relapse whereas depression symptomatology is more predictive of relapse in women (Oliva et al., 2018). Men may need additional in-home mental health supports in order to develop adaptive coping strategies for processing anxiety symptoms, especially when anxiety is comorbid with PTSD. For women, social support as part of a gender-responsive recovery context (e.g., gender-responsive recovery residence) may provide one mechanism for improving women's ability to cope with negative affect (e.g., stress, depression) during recovery.

4.3.2 Informing Gender-Responsive Sober Living Homes

Gender-specific research suggests that an empowering and supportive social context is one mechanism for promoting adaptive coping in women. This finding is grounded in women's psychology theory. Traditional theories of psychological development (e.g., stages of moral development; Kohlberg & Hersh, 1977) made basic assumptions (i.e., individuality, separation) that often classified women's relational qualities as signs of deficits rather than strengths. Jean Baker Miller (1976) and Carol Gilligan (1982) offered a new perspective on women's psychological development that framed women's relational qualities not only as strengths but as central organizing features of women's psychology. Drawing on their work, theorists developed a new theory of women's psychology, called relational-cultural theory (RCT), which has since been used to inform the development of gender-responsive treatment programming in the substance use field (Covington & Bloom, 2017). The three core tenets of RCT are: 1) the role of cultural context and its powerful effect on women's lives, 2) relationships as the central organizing feature in women's development (i.e., connection rather than separation), and 3) pathways to growth, which acknowledge women's relational qualities and activities as strengths that provide pathways to healthy growth and development (Center for Substance Abuse Treatment, 2009).

Gender-responsive programming stems from relational-cultural theory and is based on the following guiding principles (see Table 4.1 for descriptions of the principles): Gender, Environment, Relationships, Services, Socioeconomic Status, and Community (Covington & Bloom, 2017).

Table 4.1 Guiding Principles of Gender-Responsive Programming and Recommendations for Sober Living Homes

Guiding Principle	**Description**
Gender	Program acknowledges that gender makes a difference.
Environment	Program cultivates an environment based on safety, respect, and dignity.
Relationships	Program centers connection and relationships rather than separation.
Services	Program addresses substance use and mental health by offering services that are trauma-informed and inclusive (with respect to social and cultural identities).
Socioeconomic Status	Program provides women with opportunities to improve their socioeconomic conditions.
Community	Program establishes an integrated, collaborative, and comprehensive system of recovery-oriented services (e.g., Mental health providers; Alcohol and other drug treatment programs; Programs for survivors of physical and sexual violence; Family service agencies; Emergency shelter, food, and financial assistance programs; Educational organizations; Vocational training and employment services; Health care; Child welfare system, childcare, and other children's services; Transportation; Self-help groups).

Note. Adapted from Covington & Bloom (2017).

Findings from previous research and the current study together show that both women and men have gender-specific needs and may benefit from gender-responsive services. Although it is beyond the scope of the current study to make specific, concrete clinical recommendations in this area, findings indicate a need for future research and evaluations to identify whether and in what ways the services offered by sober living homes align with Covington & Bloom's (2017) guiding principles for gender-responsive services, and to identify ways in which these principles might be incorporated into sober living homes.

4.4 Conclusion

Current findings demonstrate that while social support is a strong predictor of both women's and men's recovery status during early recovery, women face unique barriers related to economic strain and psychological functioning, pointing to a need for gender-responsive services in community-based recovery contexts. Because sober living homes are organized around the idea of community, they represent an ideal setting for the implementation of gender-responsive services. It is recommended that existing sober living home programming build out gender-responsive services using established frameworks as guides (e.g., Community Assessment Tool for Gender-Responsive Programming; Covington & Bloom, 2017). Sober living homes continue to show great promise not only as a mechanism of stable housing and sober-specific social support for women and men in recovery, but as potential gender-responsive contexts that can serve to correct the health inequities that women in recovery continue to face today.

REFERENCES

Aase, D. M., Jason, L. A., Ferrari, J. R., Li, Y., & Scott, G. (2014). Comorbid mental health and substance abuse issues among individuals in recovery homes: Prospective environmental mediators. *Mental Health and Substance Use: Dual Diagnosis*, *7*(2), 170–183. https://doi.org/10.1080/17523281.2013.806342

Aldana, S. G., & Liljenquist, W. (1998). Validity and reliability of a financial strain survey. *Association for Financial Counseling and Planning Education*, *9*(2), 11–19.

Barrera, M., & Ainlay, S. L. (1983). The structure of social support: A Conceptual and empirical analysis. *Journal of Community Psychology*, *11*(2), 133–143. https://doi.org/10.1002/1520-6629(198304)11:2<133::AID-JCOP2290110207>3.0.CO;2-L

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238. https://doi.org/10.1037/0033-2909.107.2.238

Bentler, P. M. (1995). *EQS structural equations program manual* (Vol. 6). Multivariate software Encino, CA.

Bishop, P. D., Jason, L. A., & Ferrari, J. R. (1998). A survival analysis of communal-living, self-help, addiction recovery participants. *American Journal of Community Psychology*, *26*(6), 803–821. https://doi.org/10.1023/a:1022241712065

Björgvinsson, T., Kertz, S. J., Bigda-Peyton, J. S., McCoy, K. L., & Aderka, I. M. (2013). Psychometric properties of the CES-D-10 in a psychiatric sample. *Assessment*, *20*(4), 429–436. https://doi.org/10.1177/1073191113481998

Bloom, B. E., Owen, B., & Covington, S. S. (2003). *Gender-responsive strategies: Research, practice, and guiding principles for women offenders* (No. 018017; p. 142). National Institute of Corrections (NIC). https://s3.amazonaws.com/static.nicic.gov/Library/018017.pdf

Bootsmiller, B. J., Ribisl, K. M., Mowbray, C. T., Davidson, W. S., Walton, M. A., & Herman, S. E. (1998). Methods of ensuring high follow-up rates: Lessons from a longitudinal study of dual diagnosed participants. *Substance Use & Misuse*, *33*(13), 2665–2685. https://doi.org/10.3109/10826089809059344

Bowleg, L. (2012). The problem with the phrase women and minorities: Intersectionality—an important theoretical framework for public health. *American Journal of Public Health*, *102*(7), 1267–1273. https://doi.org/10.2105/AJPH.2012.300750

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Breiman, L. (2004). *Consistency for a simple model of random forests* (Technical Report 670). University of California at Berkeley. https://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=5983673EC7A3B34664FB72C813F2EC8C?doi=10.1.1.618.90&rep=rep1&type=pdf

Brem, M. J., Shorey, R. C., Anderson, S., & Stuart, G. L. (2017). Depression, anxiety, and compulsive sexual behaviour among men in residential treatment for substance use disorders: The role of experiential avoidance. *Clinical Psychology & Psychotherapy*, *24*(6), 1246–1253. https://doi.org/10.1002/cpp.2085

Brereton, K. L., Alvarez, J., Jason, L. A., Stevens, E. B., Dyson, V. B., McNeilly, C., & Ferrari, J. R. (2014). Reciprocal responsibility and social support among women in substance use recovery. *International Journal of Self Help & Self Care*, *8*(2), 239–257. https://doi.org/10.2190/SH.8.2.f

Brooks, A. T., Magaña Lòpez, M., Ranucci, A., Krumlauf, M., & Wallen, G. R. (2017). A qualitative exploration of social support during treatment for severe alcohol use disorder and recovery. *Addictive Behaviors Reports*, *6*, 76–82. https://doi.org/10.1016/j.abrep.2017.08.002

Brown, K. W., & Ryan, R. M. (2003). The benefits of being present: Mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology*, *8*(4), 822–848. https://doi.org/10.1037/0022-3514.84.4.822

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, *105*(3), 456–466. https://doi.org/10.1037/0033-2909.105.3.456

Cano, I., Best, D., Edwards, M., & Lehman, J. (2017). Recovery capital pathways: Modelling the components of recovery wellbeing. *Drug and Alcohol Dependence*, *181*, 11–19. https://doi.org/10.1016/j.drugalcdep.2017.09.002

Carlson, L. E., & Brown, K. W. (2005). Validation of the Mindful Attention Awareness Scale in a cancer population. *Journal of Psychosomatic Research*, *58*(1), 29–33. https://doi.org/10.1016/j.jpsychores.2004.04.366

Center for Substance Abuse Treatment. (2009). *Substance abuse treatment for women* (No. 51; Treatment Improvement Protocol (TIP) Series, pp. 137–180). Substance Abuse and Mental Health Services Administration (US). https://www.ncbi.nlm.nih.gov/books/NBK83257/

Centers for Disease Control and Prevention. (2020, May 20). *Health equity*. National Center for Chronic Disease Prevention and Health Promotion (NCCDPHP). https://www.cdc.gov/chronicdisease/healthequity/index.htm

Chavira, D., & Jason, L. A. (2021). 'Maybe this is the place for me:' a qualitative study examining needs and length of stay in recovery house residents. *Journal of Social Work Practice in the Addictions*, *0*(0), 1–12. https://doi.org/10.1080/1533256X.2021.1916174

Cloud, W., & Granfield, R. (2008). Conceptualizing recovery capital: Expansion of a theoretical construct. *Substance Use & Misuse*, *43*(12–13), 1971–1986. https://doi.org/10.1080/10826080802289762

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, 385–396. https://doi.org/10.2307/2136404

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Academic Press.

Covington, S. S. (2002). Helping women recover: Creating gender-responsive treatment. In S. L. A. Straussner & S. Brown (Eds.), *The handbook of addiction treatment for women: Theory and practice* (p. 18). Jossey-Bass.

Covington, S. S., & Bloom, B. E. (2017). *Gender-responsive program assessment tool*. Institute for Relational Development. https://stephaniecovington.com/site/assets/files/1303/5796-65-genderresponsiveassessment_community.pdf

Digiusto, E., Panjari, M., Gibson, A., & Rea, F. (2006). Follow-up difficulty: Correlates and relationship with outcome in heroin dependence treatment in the NEPOD study. *Addictive Behaviors*, 1201–1210. https://doi.org/10.1016/j.addbeh.2005.09.004

Drapeau, A., Beaulieu-Prévost, D., Marchand, A., Boyer, R., Préville, M., & Kairouz, S. (2010). A life-course and time perspective on the construct validity of psychological distress in women and men: Measurement invariance of the K6 across gender. *BMC Medical Research Methodology*, *10*, 68. https://doi.org/10.1186/1471-2288-10-68

Edwards, K. M., Murphy, S., Palmer, K. M., Haynes, E. E., Chapo, S., Ekdahl, B. A., & Buel, S. (2017). Co-occurrence of and recovery from substance abuse and lifespan victimization: A qualitative study of female residents in trauma-informed sober living homes. *Journal of Psychoactive Drugs*, *49*(1), 74–82. https://doi.org/10.1080/02791072.2016.1273566

Edwards, K. M., Siller, L., Ullman, S. E., Lee, K. D. M., & Murphy, S. B. (2021). Post-traumatic growth in women with histories of addiction and victimization residing in a sober living home. *Journal of Interpersonal Violence,* 1–18. https://doi.org/10.1177/0886260521991283

Edwards, K. M., Ullman, S. E., Siller, L., Murphy, S. B., Harvey, R., & Wheeler, L. (2022). What relates to length of stay in a sober living home among women with histories of victimization and addiction? A longitudinal analysis. *Journal of Community Psychology*. https://doi.org/10.1002/jcop.22792

Emmanuel, T., Maupong, T., Mpoeleng, D., Semong, T., Mphago, B., & Tabona, O. (2021). A survey on missing data in machine learning. *Journal of Big Data*, *8*(1), 140. https://doi.org/10.1186/s40537-021-00516-9

Enders, C. K. (2001). A primer on maximum likelihood algorithms available for use with missing data. *Structural Equation Modeling: A Multidisciplinary Journal*, *8*(1), 128–141. https://doi.org/10.1207/S15328007SEM0801_7

Enders, C. K. (2011). Missing not at random models for latent growth curve analyses. *Psychological Methods*, *16*(1), 1. https://doi.org/10.1037/a0022640

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *8*(3), 430–457. https://doi.org/10.1207/S15328007SEM0803_5

Foster, K. T., Li, N., McClure, E. A., Sonne, S. C., & Gray, K. M. (2016). Gender differences in internalizing symptoms and suicide risk among men and women seeking treatment for cannabis use disorder from late adolescence to middle adulthood. *Journal of Substance Abuse Treatment*, *66*, 16–22. https://doi.org/10.1016/j.jsat.2016.01.012

Gilligan, C. (1982). *In a different voice*. Harvard University Press.

Granfield, R., & Cloud, W. (1999). *Coming clean: Overcoming addiction without treatment*. NYU Press. https://nyupress.org/9780814715826/coming-clean

Greenfield, S. F., Brooks, A. J., Gordon, S. M., Green, C. A., Kropp, F., McHugh, R. K., Lincoln, M., Hien, D., & Miele, G. M. (2007). Substance abuse treatment entry, retention, and outcome in women: A review of the literature. *Drug and*

Alcohol Dependence, *86*(1), 1–21. https://doi.org/10.1016/j.drugalcdep.2006.05.012

Greenfield, S. F., & Grella, C. E. (2009). Alcohol & drug abuse: What is "women-focused" treatment for substance use disorders? *Psychiatric Services*, *60*(7), 880–882. https://doi.org/10.1176/ps.2009.60.7.880

Grella, C. E. (2008). From generic to gender-responsive treatment: Changes in social policies, treatment services, and outcomes of women in substance abuse treatment. *Journal of Psychoactive Drugs*, *40*(sup5), 327–343. https://doi.org/10.1080/02791072.2008.10400661

Grella, C. E., Scott, C. K., & Foss, M. A. (2005). Gender differences in long-term drug treatment outcomes in Chicago PETS. *Journal of Substance Abuse Treatment*, *28*(2, Supplement), S3–S12. https://doi.org/10.1016/j.jsat.2004.08.008

Hallgren, K. A., McCabe, C. J., King, K. M., & Atkins, D. C. (2019). Beyond path diagrams: Enhancing applied structural equation modeling research through data visualization. *Addictive Behaviors*, *94*, 74–82. https://doi.org/10.1016/j.addbeh.2018.08.030

Hansten, M. L., Downey, L., Rosengren, D. B., & Donovan, D. M. (2000). Relationship between follow-up rates and treatment outcomes in substance abuse research: More is better but when is "enough" enough? *Addiction (Abingdon, England)*, *95*(9), 1403–1416. https://doi.org/10.1046/j.1360-0443.2000.959140310.x

Harvey, R., Jason, L. A., & Ferrari, J. R. (2016). Substance abuse relapse in Oxford House recovery homes: A survival analysis evaluation. *Substance Abuse*, *37*(2), 281–285. https://doi.org/10.1080/08897077.2015.1080786

Haseltine, F. P. (2000). Gender differences in addiction and recovery. *Journal of Women's Health & Gender-Based Medicine*, *9*(6), 579–583. https://doi.org/10.1089/15246090050118080

Hasin, D. S., O'Brien, C. P., Auriacombe, M., Borges, G., Bucholz, K., Budney, A., Compton, W. M., Crowley, T., Ling, W., Petry, N. M., Schuckit, M., & Grant, B. F. (2013). DSM-5 criteria for substance use disorders: Recommendations and rationale. *The American Journal of Psychiatry*, *170*(8), 834. https://doi.org/10.1176/appi.ajp.2013.12060782

Hastie, T., Tibshirani, R., & Friedman, J. (2009). Overview of Supervised Learning. In T. Hastie, R. Tibshirani, & J. Friedman, *The Elements of Statistical Learning* (pp. 9–41). Springer New York. https://doi.org/10.1007/978-0-387-84858-7_2

Hayes, S. C., Wilson, K. G., Gifford, E. V., Follette, V. M., & Strosahl, K. (1996). Experimental avoidance and behavioral disorders: A functional dimensional approach to diagnosis and treatment. *Journal of Consulting and Clinical Psychology*, *64*(6), 1152–1168. https://doi.org/10.1037//0022-006x.64.6.1152

Hayward, M. D., Miles, T. P., Crimmins, E. M., & Yang, Y. (2000). The significance of socioeconomic status in explaining the racial gap in chronic health conditions. *American Sociological Review*, *65*(6), 910–930. https://doi.org/10.2307/2657519

Hecksher, D., & Hesse, M. (2009). Women and substance use disorders. *Mens Sana Monographs*, *7*(1), 50–62. https://doi.org/10.4103/0973-1229.42585

Hernandez-Avila, C. A., Rounsaville, B. J., & Kranzler, H. R. (2004). Opioid-, cannabis- and alcohol-dependent women show more rapid progression to substance abuse treatment. *Drug and Alcohol Dependence*, *74*(3), 265–272. https://doi.org/10.1016/j.drugalcdep.2004.02.001

Heron, J., Hickman, M., Macleod, J., & Munafò, M. R. (2011). Characterizing patterns of smoking initiation in adolescence: Comparison of methods for dealing with missing data. *Nicotine & Tobacco Research*, *13*(12), 1266–1275. https://doi.org/10.1093/ntr/ntr161

Jason, L. A., Davis, M. I., & Ferrari, J. R. (2007). The need for substance abuse after-care: Longitudinal analysis of Oxford House. *Addictive Behaviors*, *32*(4), 803–818. https://doi.org/10.1016/j.addbeh.2006.06.014

Jason, L. A., Salina, D., & Ram, D. (2016). Oxford recovery housing: Length of stay correlated with improved outcomes for women previously involved with the criminal justice system. *Substance Abuse*, *37*(1), 248–254. https://doi.org/10.1080/08897077.2015.1037946

Jason, L. A., Wiedbusch, E., Bobak, T. J., & Taullahu, D. (2020). Estimating the number of substance use disorder recovery homes in the United States. *Alcoholism Treatment Quarterly*, *38*(4), 506–514. https://doi.org/10.1080/07347324.2020.1760756

Kahn, J. R., & Pearlin, L. I. (2006). Financial strain over the life course and health among older adults. *Journal of Health and Social Behavior*, *47*(1), 17–31. https://doi.org/10.1177/002214650604700102

Karpyak, V. M., Biernacka, J. M., Geske, J. R., Abulseoud, O. A., Brunner, M. D., Chauhan, M., Hall-Flavin, D. K., Lewis, K. A., Loukianova, L. L., Melnyk, G. J., Onsrud, D. A., Proctor, B. D., Schneekloth, T. D., Skime, M. K., Wittkopp, J. E., Frye, M. A., & Mrazek, D. A. (2016). Gender-specific effects of comorbid depression and anxiety on the propensity to drink in negative emotional states. *Addiction (Abingdon, England)*, *111*(8), 1366–1375. https://doi.org/10.1111/add.13386

Kelly, B. C., Carpiano, R. M., Easterbrook, A., & Parsons, J. T. (2014). Exploring the gay community question: Neighborhood and network influences on the experience of community among urban gay men. *The Sociological Quarterly*, *55*(1), 23–48. https://doi.org/10.1111/tsq.12041

Kohlberg, L., & Hersh, R. H. (1977). Moral development: A review of the theory. *Theory Into Practice*, *16*(2), 53–59. https://doi.org/10.1080/00405847709542675

Krentzman, A. R., Sharratt, M., Mannella, K. A., Foster, K. T., Yarosh, S., & Jester, J. M. (2022). Women in sober living: Baseline characteristics and their associations with reasons for leaving. *Journal of Social Work Practice in the Addictions*, *22*(1), 21–39. https://doi.org/10.1080/1533256X.2021.1935156

Laudet, A. B., Cleland, C. M., Magura, S., Vogel, H. S., & Knight, E. L. (2004). Social support mediates the effects of dual-focus mutual aid groups on abstinence from substance use. *American Journal of Community Psychology*, *34*(3–4), 175–185. https://doi.org/10.1007/s10464-004-7413-5

Laudet, A. B., Savage, R., & Mahmood, D. (2002). Pathways to long-term recovery: A preliminary investigation. *Journal of Psychoactive Drugs*, *34*(3), 305–311. https://doi.org/10.1080/02791072.2002.10399968

Laudet, A. B., & White, W. L. (2008). Recovery capital as prospective predictor of sustained recovery, life satisfaction, and stress among former poly-substance users. *Substance Use & Misuse*, *43*(1), 27–54. https://doi.org/10.1080/10826080701681473

Liu, X., & Kaplan, H. B. (1996). Gender-related differences in circumstances surrounding initiation and escalation of alcohol and other substance use/abuse. *Deviant Behavior*, *17*(1), 71–106. https://doi.org/10.1080/01639625.1996.9968015

Logan, T., Walker, R., Jordan, C. E., & Leukefeld, C. G. (2006). *Women and victimization: Contributing factors, interventions, and implications* (pp. xii, 321). American Psychological Association. https://doi.org/10.1037/11364-000

Lotzin, A., Buth, S., Sehner, S., Hiller, P., Martens, M.-S., Read, J., Härter, M., Cowlishaw, S., & Schäfer, I. (2019). Learning How to Ask - Does a one-day training increase trauma inquiry in routine substance use disorder practice?

Results of a cluster-randomized controlled trial. *Journal of Substance Abuse Treatment*, *107*, 8–16. https://doi.org/10.1016/j.jsat.2019.08.005

Loveland, D., & Driscoll, H. (2014). Examining attrition rates at one specialty addiction treatment provider in the United States: A case study using a retrospective chart review. *Substance Abuse Treatment, Prevention, and Policy*, *9*(1), 41. https://doi.org/10.1186/1747-597X-9-41

Mahoney, E., Witbrodt, J., Mericle, A. A., & Polcin, D. L. (2021). Resident and house manager perceptions of social environments in sober living houses: Associations with length of stay. *Journal of Community Psychology*, *49*(7), 2959–2971. https://doi.org/10.1002/jcop.22620

Marcoulides, K. M., & Yuan, K.-H. (2017). New ways to evaluate goodness of fit: A note on using equivalence testing to assess structural equation models. *Structural Equation Modeling*, *24*(1), 148–153. https://doi.org/10.1080/10705511.2016.1225260

Mason, M. J. (1999). A review of procedural and statistical methods for handling attrition and missing data in clinical research. *Measurement and Evaluation in Counseling and Development*, *32*(2), 111–118. https://doi.org/10.1080/07481756.1999.12068976

McGaffin, B. J., Deane, F. P., Kelly, P. J., & Blackman, R. J. (2018). Social support and mental health during recovery from drug and alcohol problems. *Addiction Research & Theory*, *26*(5), 386–395. https://doi.org/10.1080/16066359.2017.1421178

McHugh, R. K., DeVito, E. E., Dodd, D., Carroll, K. M., Potter, J. S., Greenfield, S. F., Connery, H. S., & Weiss, R. D. (2013). Gender differences in a clinical trial for prescription opioid dependence. *Journal of Substance Abuse Treatment*, *45*(1), 38–43. https://doi.org/10.1016/j.jsat.2012.12.007

McHugh, R. K., Votaw, V. R., Sugarman, D. E., & Greenfield, S. F. (2018). Sex and gender differences in substance use disorders. *Clinical Psychology Review*, *66*, 12–23. https://doi.org/10.1016/j.cpr.2017.10.012

Mericle, A. A., Carrico, A. W., Hemberg, J., Stall, R., & Polcin, D. L. (2019). Improving recovery outcomes among MSM: The potential role of recovery housing. *Journal of Substance Use*, *24*(2), 140–146. https://doi.org/10.1080/14659891.2018.1523966

Meyer, I. H. (2003). Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: Conceptual issues and research evidence. *Psychological Bulletin*, *129*(5), 674–697. https://doi.org/10.1037/0033-2909.129.5.674

Miller, J. B. (1976). *Toward a new psychology of women*. Beacon Press.

Mohammed, M. B., Zulkafli, H. S., Adam, M. B., Ali, N., & Baba, I. A. (2021). Comparison of five imputation methods in handling missing data in a continuous frequency table. *AIP Conference Proceedings*, *2355*(1), 040006. https://doi.org/10.1063/5.0053286

Muthén, B., Asparouhov, T., Hunter, A. M., & Leuchter, A. F. (2011). Growth modeling with nonignorable dropout: Alternative analyses of the STAR*D antidepressant trial. *Psychological Methods*, *16*(1), 17. https://doi.org/10.1037/a0022634

National Council for Behavioral Health, & National Alliance for Recovery Residences. (2018). *Building recovery: State policy guide for supporting recovery housing* (p. 42) [Toolkit]. National Council for Behavioral Health. https://www.thenationalcouncil.org/wp-content/uploads/2018/05/18_Recovery-Housing-Toolkit_5.3.2018.pdf?daf=375ateTbd56

National Institute of Mental Health. (2020, January). *Bipolar Disorder*. National Institute of Mental Health (NIMH). https://www.nimh.nih.gov/health/topics/bipolar-disorder

National Institute on Drug Abuse. (2020, June 3). *How long does drug addiction treatment usually last?* National Institute on Drug Abuse. https://nida.nih.gov/publications/principles-drug-addiction-treatment-research-based-guide-third-edition/frequently-asked-questions/how-long-does-drug-addiction-treatment-usually-last

Nealon-Woods, M. A., Ferrari, J. R., & Jason, L. A. (1995). Twelve-step program use among Oxford house residents: Spirituality or social support in sobriety? *Journal of Substance Abuse*, *7*(3), 311–318. https://doi.org/10.1016/0899-3289(95)90024-1

Nemes, S., Wish, E., Wraight, B., & Messina, N. (2002). Correlates of treatment follow-up difficulty. *Substance Use & Misuse*, *37*(1), 19–45. https://doi.org/10.1081/JA-120001495

Oliva, F., Nibbio, G., Vizzuso, P., Jaretti Sodano, A., Ostacoli, L., Carletto, S., & Picci, R. L. (2018). Gender differences in anxiety and depression before and after alcohol detoxification: Anxiety and depression as gender-related predictors of relapse. *European Addiction Research*, *24*, 163–172. https://doi.org/10.1159/000490046

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., & Cournapeau, D. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830. https://arxiv.org/abs/1201.0490

Peugh, J. L., & Enders, C. K. (2004). Missing data in educational research: A review of reporting practices and suggestions for improvement. *Review of Educational Research*, *74*(4), 525–556. https://doi.org/10.3102/00346543074004525

Pirneskoski, J., Tamminen, J., Kallonen, A., Nurmi, J., Kuisma, M., Olkkola, K. T., & Hoppu, S. (2020). Random forest machine learning method outperforms prehospital National Early Warning Score for predicting one-day mortality: A retrospective study. *Resuscitation Plus*, *4*, 100046. https://doi.org/10.1016/j.resplu.2020.100046

Polcin, D. L. (2001). Sober living houses: Potential roles in substance abuse services and suggestions for research. *Substance Use & Misuse*, *36*(3), 301–311. https://doi.org/10.1081/JA-100102627

Polcin, D. L., & Borkman, T. (2008). The impact of AA on non-professional substance abuse recovery programs and sober living houses. *Recent Developments in Alcoholism: An Official Publication of the American Medical Society on Alcoholism, the Research Society on Alcoholism, and the National Council on Alcoholism*, *18*, 91–108. https://doi.org/10.1007/978-0-387-77725-2_6

Polcin, D. L., & Korcha, R. (2017). Social support influences on substance abuse outcomes among sober living house residents with low and moderate psychiatric severity. *Journal of Alcohol and Drug Education*, *61*(1), 51–70. https://www.ncbi.nlm.nih.gov/pubmed/28757663

Polcin, D. L., Korcha, R. A., Bond, J., & Galloway, G. (2010a). Sober living houses for alcohol and drug dependence: 18-Month outcomes. *Journal of Substance Abuse Treatment*, *38*(4), 356–365. https://doi.org/10.1016/j.jsat.2010.02.003

Polcin, D. L., Korcha, R., Bond, J., & Galloway, G. (2010). What did we learn from our study on sober living houses and where do we go from here? *Journal of Psychoactive Drugs*, *42*(4), 425–433. https://dx.doi.org/10.1080%2F02791072.2010.10400705

Polich, J. M., Armor, D. J., & Braiker, H. B. (1980). *The course of alcoholism: Four years after treatment* (No. ADA484014). Rand Corporation. https://apps.dtic.mil/sti/citations/ADA484014

Prendergast, M. L., Messina, N. P., Hall, E. A., & Warda, U. S. (2011). The relative effectiveness of women-only and mixed-gender treatment for substance-abusing women. *Journal of Substance Abuse Treatment*, *40*(4), 336–348. https://doi.org/10.1016/j.jsat.2010.12.001

Qi, Y. (2012). Random forest for bioinformatics. In *Ensemble machine learning* (pp. 307–323). Springer.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*(3), 385–401. https://doi.org/10.1177/014662167700100306

Raistrick, D., Bradshaw, J., Tober, G., Weiner, J., Allison, J., & Healey, C. (1994). Development of the Leeds Dependence Questionnaire (LDQ): A questionnaire to measure alcohol and opiate dependence in the context of a treatment evaluation package. *Addiction (Abingdon, England)*, *89*(5), 563–572. https://doi.org/10.1111/j.1360-0443.1994.tb03332.x

Ramchandani, P. (2018, October 10). *Random Forests and the Bias-Variance Tradeoff*. Medium. https://towardsdatascience.com/random-forests-and-the-bias-variance-tradeoff-3b77fee339b4

Recovery Research Institute. (2022). *Addictionary—Glossary of Substance Use Disorder Terminology*. https://www.recoveryanswers.org/addiction-ary/

Rhodes, D. R., & Johnson, M. A. (1997). A feminist approach to treating alcohol and drug-addicted African-American women. *Women & Therapy*, *20*(3), 23–37. https://doi.org/10.1300/J015v20n03_02

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*, 1–36. https://doi.org/10.18637/jss.v048.i02

Sass, D. A., & Schmitt, T. A. (2013). Testing measurement and structural invariance. In T. Teo (Ed.), *Handbook of Quantitative Methods for Educational Research* (pp. 315–345). SensePublishers. https://doi.org/10.1007/978-94-6209-404-8_15

Scott, C. K. (2004). A replicable model for achieving over 90% follow-up rates in longitudinal studies of substance abusers. *Drug and Alcohol Dependence*, *74*(1), 21–36. https://doi.org/10.1016/j.drugalcdep.2003.11.007

Sherbourne, C. D., & Stewart, A. L. (1991). The MOS social support survey. *Social Science & Medicine*, *32*(6), 705–714. https://doi.org/10.1016/0277-9536(91)90150-B

Shippee, T. P., Wilkinson, L. R., & Ferraro, K. F. (2012). Accumulated financial strain and women's health over three decades. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *67*(5), 585–594. https://doi.org/10.1093/geronb/gbs056

Smith, J. A., Franklin, S., Asikis, C., Knudsen, S., Woodruff, A., & Kimball, T. (2018). Social support and gender as correlates of relapse risk in collegiate recovery programs. *Alcoholism Treatment Quarterly*, *36*(3), 354–365. https://doi.org/10.1080/07347324.2018.1437372

Smith, L. R., Earnshaw, V. A., Copenhaver, M. M., & Cunningham, C. O. (2016). Substance use stigma: Reliability and validity of a theory-based scale for substance-using populations. *Drug and Alcohol Dependence*, *162*, 34–43. https://doi.org/10.1016/j.drugalcdep.2016.02.019

Steiger, J. H., & Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, *25*(2), 173–180. https://doi.org/10.1207/s15327906mbr2502_4

Stevens, E., Jason, L. A., Ram, D., & Light, J. (2015). Investigating social support and network relationships in substance use disorder recovery. *Substance Abuse*, *36*(4), 396–399. https://doi.org/10.1080/08897077.2014.965870

Stevens, L. M., Mortazavi, B. J., Deo, R. C., Curtis, L., & Kao, D. P. (2020). Recommendations for reporting machine learning analyses in clinical research. *Circulation: Cardiovascular Quality and Outcomes*, *13*(10), e006556. https://doi.org/10.1161/CIRCOUTCOMES.120.006556

Substance Abuse and Mental Health Services Administration, Office of the Surgeon General (US). (2016). *Facing addiction in America: The surgeon general's report on alcohol, drugs, and health*. U.S. Department of Health and Human Services. https://addiction.surgeongeneral.gov/sites/default/files/surgeon-generals-report.pdf

Svetnik, V., Liaw, A., Tong, C., Culberson, J. C., Sheridan, R. P., & Feuston, B. P. (2003). Random forest: A classification and regression tool for compound classification and QSAR modeling. *Journal of Chemical Information and Computer Sciences*, *43*(6), 1947–1958. https://doi.org/10.1021/ci034160g

Timko, C., Finney, J. W., & Moos, R. H. (2005). The 8-year course of alcohol abuse: Gender differences in social context and coping. *Alcoholism: Clinical and Experimental Research*, *29*(4), 612–621. https://doi.org/0145-6008/05/2904-0612$03.00/0

Toomey, R. B., Ryan, C., Diaz, R. M., & Russell, S. T. (2018). Coping with sexual orientation-related minority stress. *Journal of Homosexuality*, *65*(4), 484–500. https://doi.org/10.1080/00918369.2017.1321888

Tracy, E. M., Munson, M. R., Peterson, L. T., & Floersch, J. E. (2010). Social support: A mixed blessing for women in substance abuse treatment. *Journal of Social Work Practice in the Addictions*, *10*(3), 257–282. https://doi.org/10.1080/1533256X.2010.500970

Tuchman, E. (2011). Women and addiction: The importance of gender issues in substance abuse research. *FOCUS*, *9*(1), 90–90. https://doi.org/10.1176/foc.9.1.foce90

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*(1), 1–10. https://doi.org/10.1007/BF02291170

U.S. Bureau of Labor Statistics. (2021, September 16). *Labor force statistics from the Current Population Survey*. https://www.bls.gov/cps/earnings.htm

Wadekar, A. S. (2020). Understanding opioid use disorder (OUD) using tree-based classifiers. *Drug and Alcohol Dependence*, *208*, 107839. https://doi.org/10.1016/j.drugalcdep.2020.107839

Weisberg, H. F. (2005). *The total survey error approach: A guide to the new science of survey research*. University of Chicago Press. https://press.uchicago.edu/ucp/books/book/chicago/T/bo3619292.html

Wu, L.-T., Ling, W., Burchett, B., Blazer, D. G., Shostak, J., & Woody, G. E. (2010). Gender and racial/ethnic differences in addiction severity, HIV risk, and quality of life among adults in opioid detoxification: Results from the National Drug Abuse Treatment Clinical Trials Network. *Substance Abuse and Rehabilitation*, *1*, 13–22. https://doi.org/10.2147/SAR.S15151

Yuan, K.-H., Chan, W., Marcoulides, G. A., & Bentler, P. M. (2016). Assessing structural equation models by equivalence testing with adjusted fit indexes. *Structural Equation Modeling*, *23*(3), 319–330. https://doi.org/10.1080/10705511.2015.1065414

Appendix A

EXAMPLE OF MISSING DATA PATTERNS AND INITIAL CATEGORIES FOR SOCIAL SUPPORT

% Missing	Time 1	Time 2	Time 3	Time 4	Time 5	Time 6	Missing Data Category
38%	1	1	1	1	1	1	Complete
13%	1	0	0	0	0	0	Lost to Follow Up
10%	1	1	0	0	0	0	Lost to Follow Up
8%	1	1	1	0	0	0	Lost to Follow Up
6%	1	1	1	1	0	0	Lost to Follow Up
6%	1	1	1	1	1	0	Sporadic
2%	1	1	0	1	1	1	Sporadic
2%	0	1	0	0	0	0	One Wave Only
2%	1	0	1	0	0	0	Lost to Follow Up
2%	1	0	1	1	0	1	Sporadic
2%	1	0	1	1	1	1	Sporadic
2%	1	1	0	1	0	0	Lost to Follow Up
2%	1	1	1	0	1	1	Sporadic
< 1%	0	1	1	1	1	1	Sporadic
< 1%	1	0	0	0	1	1	Sporadic
< 1%	1	0	0	1	1	0	Sporadic
< 1%	1	0	0	1	1	1	Sporadic
< 1%	1	1	0	0	1	0	Sporadic
< 1%	1	1	0	1	0	1	Sporadic
< 1%	1	1	1	0	1	0	Sporadic
< 1%	1	1	1	1	0	1	Sporadic

Appendix B

CLASSIFICATIONS, SUGGESTED CUTOFF VALUES, AND DESCRIPTIONS OF REPORTED MODEL FIT INDICES

Fit indices	**Class**	**Suggested model fit cut-off values**	**Description**
Standardized root mean-squared residual (SRMR; Bentler, 1995)	Absolute	SRMR < .08 (Hu & Bentler, 1999; Mueller & Hancock, 2018)	Standardized discrepancy between observed and predicted correlations; ranges from 0.0 to 1.0 with smaller values indicating better model fit (Wickrama et al., 2016).
Root mean square error of approximation (RMSEA; Steiger, 1990)	Parsimonious	RMSEA < .06 $p < .05$ (Hu & Bentler, 1999) RMSEA < .08 $p < .05$ (Mueller & Hancock, 2018)	Fit index that compensates for model complexity; relatively insensitive to sample size; values of 0 indicate perfect model fit to the data (Wickrama et al., 2016).
Tucker-Lewis Index (TLI; Tucker & Lewis, 1973)	Incremental	TLI ≥ .95 (Hu & Bentler, 1999)	Indicates the model of interest improves the fit by 95% relative to the null model (Byrne, 1998); preferable for smaller samples (Tabachnick and Fidell, 2007).
Comparative fit index (CFI; Bentler, 1990)	Incremental	CFI ≥ .95 (Hu & Bentler, 1999)	Evaluates the fit of a hypothesized model compared to a reduced nested model (known as the restricted model) (Hu & Bentler, 1998).

Appendix C

FEATURES USED IN FULL SAMPLE RANDOM FOREST MODEL

Sociodemographics Variables					
Variable Name in Model	Variable	# of categories/ range	Variable Type	Encoding	Description
black	Black, non-Hispanic	2	Binary Categorical	Dummy	Black, non-Hispanic vs. all other race/ethnicity categories.
white	White, non-Hispanic	2	Binary Categorical	Dummy	White, non-Hispanic vs. all other race/ethnicity categories.
single	Relationship Status: Single	2	Binary Categorical	Dummy	Single vs. Other Relationship Status Categories (i.e., Dating, Married, In a relationship, Separated, Divorced, Widowed).
educlevel	Education Level	7	Ordinal	Integer	Education Level Categories: Some High school, High School Degree, Some College, 2-yr. College Degree, 4-yr. College Degree,

					Some Grad School, Grad School Degree
age	Age (numeric)	19 - 64	Continuous	Float	Age at Baseline
lgbtq	LGBTQ status	2	Binary Categorical	Dummy	Binary indicator of LGBTQ status.
women	Gender	2	Binary Categorical	Dummy	Binary indicator of gender (women = 1) where all individuals who identified as women, including Transgender Women, were coded as 1 and all individuals who identified as men, including Transgender Men, were coded as 0.
employed	Employed (Full-time/Part-time)	2	Binary Categorical	Dummy	Binary indicator of Employment.
children	Parental Status (Has Children)	2	Binary Categorical	Dummy	Binary indicator of parental status (1=Has children)
currparole	Currently on Parole	2	Binary Categorical	Dummy	Participant was on parole at Baseline.
everjail	Previously in Jail or Incarcerated	2	Binary Categorical	Dummy	Binary indicator of jail/ incarceration experience across lifetime.
Mental Health Variables					
depression	Major Depression Disorder Diagnosis	2	Binary Categorical	Dummy	Binary indicator of Major Depression Disorder Diagnosis.

nmhd	Number of Mental Health Disorder Diagnoses	8	Ordinal	Integer	0-7
bipolar1	Bipolar I Diagnosis	2	Binary Categorical	Dummy	Binary indicator of Bipolar I Diagnosis.
anxiety	Anxiety Disorder Diagnosis	2	Binary Categorical	Dummy	Binary indicator of Anxiety Disorder Diagnosis.
ptsd	Posttraumatic Stress Disorder (PTSD) Diagnosis	2	Binary Categorical	Dummy	Binary indicator of PTSD Diagnosis.
Substance Use Variables					
daysih_max	Average Length of Stay in Home	1-1,260	Continuous	Float	Average length of stay in home (days).
lifetimerelapse	Participant experienced a relapse during lifetime	2	Binary Categorical	Dummy	Binary indicator of whether or not a participant experienced a relapse (i.e., recurrence of SUD symptoms after period of abstinence) at some point in their lifetime.
limen	Recovery Housing Organization	2	Binary Categorical	Dummy	Binary indicator of Recovery Housing Organization, where 1 = Organization 1 and 0 = Organization 2.
odfreq	Number of Lifetime Overdoses at Baseline	0-15	Continuous	Float	Number of overdoses experienced by participant across their lifetime, assessed at baseline.

movedout	Participant moved out of home at some point during study	2	Binary Categorical	Dummy	Binary indicator of participant moving out of sober living home at some point during the study.
primsub_1.0	Primary Substance of Use: Alcohol	2	Binary Categorical	Dummy	Binary indicator of Alcohol as participant’s primary substance of use.
primsub_2.0	Primary Substance of Use: Cocaine	2	Binary Categorical	Dummy	Binary indicator of Cocaine as participant’s primary substance of use.
primsub_3.0	Primary Substance of Use: Opioids	2	Binary Categorical	Dummy	Binary indicator of Opioids as participant’s primary substance of use.
primsub_4.0	Primary Substance of Use: Other	2	Binary Categorical	Dummy	Binary indicator of Other as participant’s primary substance of use.
recovdays	Number of Days in Recovery at Baseline	4-5,475	Continuous	Float	Number of days in recovery assessed at baseline.
severecat	Addiction Severity	3	Categorical	Integer	Hierarchical categorical indicator of addiction severity: 1=mild (0-20) 2=moderate (21-25) 3=severe (26-30)
Recovery Capital Features					
pred_dep1	Predicted Depression at Wave 1	3.59-18.97	Continuous	Float	Individual predicted values for depression at Baseline.

pred_dep2	Predicted Depression at Wave 2	2.75-19.63	Continuous	Float	Individual predicted values for depression at Month 2.
pred_dep3	Predicted Depression at Wave 3	1.89-20.29	Continuous	Float	Individual predicted values for depression at Month 3.
pred_dep4	Predicted Depression at Wave 4	1.04-20.95	Continuous	Float	Individual predicted values for depression at Month 4.
pred_dep5	Predicted Depression at Wave 5	0.18-21.61	Continuous	Float	Individual predicted values for depression at Month 5.
pred_dep6	Predicted Depression at Wave 6	-0.68-22.27	Continuous	Float	Individual predicted values for depression at Month 6.
pred_stress1	Predicted Perceived Stress at Wave 1	2.05-10.87	Continuous	Float	Individual predicted values for perceived stress at Baseline.
pred_stress2	Predicted Perceived Stress at Wave 2	1.61-10.64	Continuous	Float	Individual predicted values for perceived stress at Month 2.
pred_stress3	Predicted Perceived Stress at Wave 3	1.16-10.41	Continuous	Float	Individual predicted values for perceived stress at Month 3.
pred_stress4	Predicted Perceived Stress at Wave 4	0.72-10.18	Continuous	Float	Individual predicted values for perceived stress at Month 4.
pred_stress5	Predicted Perceived Stress at Wave 5	0.27-9.94	Continuous	Float	Individual predicted values for perceived stress at Month 5.
pred_stress6	Predicted Perceived Stress at Wave 6	-0.18-9.71	Continuous	Float	Individual predicted values for perceived stress at Month 6.
pred_finstrain1	Predicted Financial Strain at Wave 1	5.11-24.89	Continuous	Float	Individual predicted values for financial strain at Baseline.

pred_finstrain2	Predicted Financial Strain at Wave 2	4.87-24.79	Continuous	Float	Individual predicted values for financial strain at Month 2.
pred_finstrain3	Predicted Financial Strain at Wave 3	4.63-25.29	Continuous	Float	Individual predicted values for financial strain at Month 3.
pred_finstrain4	Predicted Financial Strain at Wave 4	4.39-25.79	Continuous	Float	Individual predicted values for financial strain at Month 4.
pred_finstrain5	Predicted Financial Strain at Wave 5	4.15-26.29	Continuous	Float	Individual predicted values for financial strain at Month 5.
pred_finstrain6	Predicted Financial Strain at Wave 6	3.91-26.79	Continuous	Float	Individual predicted values for financial strain at Month 6.
pred_socsupp1	Predicted Social Support at Wave 1	0.78-4.06	Continuous	Float	Individual predicted values for social support at Baseline.
pred_socsupp2	Predicted Social Support at Wave 2	1.17-3.8	Continuous	Float	Individual predicted values for social support at Month 2.
pred_socsupp3	Predicted Social Support at Wave 3	1.31-3.84	Continuous	Float	Individual predicted values for social support at Month 3.
pred_socsupp4	Predicted Social Support at Wave 4	1.26-3.91	Continuous	Float	Individual predicted values for social support at Month 4.
pred_socsupp5	Predicted Social Support at Wave 5	0.77-3.99	Continuous	Float	Individual predicted values for social support at Month 5.
pred_socsupp6	Predicted Social Support at Wave 6	0.14-4.2	Continuous	Float	Individual predicted values for social support at Month 6.
resid_dep1	Residual Depression at Wave 1	-9.56-10.55	Continuous	Float	Residual values for Depression at Baseline calculated as the difference between the observed values and the

					predicted values.
resid_dep2	Residual Depression at Wave 2	-12.24-18.5	Continuous	Float	Residual values for Depression at Month 2 calculated as the difference between the observed values and the predicted values.
resid_dep3	Residual Depression at Wave 3	-11.26-12.25	Continuous	Float	Residual values for Depression at Month 3 calculated as the difference between the observed values and the predicted values.
resid_dep4	Residual Depression at Wave 4	-11.61-12.71	Continuous	Float	Residual values for Depression at Month 4 calculated as the difference between the observed values and the predicted values.
resid_dep5	Residual Depression at Wave 5	-8.52-11.99	Continuous	Float	Residual values for Depression at Month 5 calculated as the difference between the observed values and the predicted values.
resid_dep6	Residual Depression at Wave 6	-15.53-9.78	Continuous	Float	Residual values for Depression at Month 6 calculated as the difference between the observed values and the predicted values.
resid_stress1	Residual Perceived Stress at Wave 1	-4.37-5.47	Continuous	Float	Residual values for Perceived stress at Baseline calculated as the difference between the observed values and the predicted values.
resid_stress2	Residual Perceived Stress	-4.77-6.61	Continuous	Float	Residual values for Perceived stress

	at Wave 2				at Month 2 calculated as the difference between the observed values and the predicted values.
resid_stress3	Residual Perceived Stress at Wave 3	-3.88-5.44	Continuous	Float	Residual values for Perceived stress at Month 3 calculated as the difference between the observed values and the predicted values.
resid_stress4	Residual Perceived Stress at Wave 4	-5.6-6.45	Continuous	Float	Residual values for Perceived stress at Month 4 calculated as the difference between the observed values and the predicted values.
resid_stress5	Residual Perceived Stress at Wave 5	-3.39-6.96	Continuous	Float	Residual values for Perceived stress at Month 5 calculated as the difference between the observed values and the predicted values.
resid_stress6	Residual Perceived Stress at Wave 6	-3.66-4.04	Continuous	Float	Residual values for Perceived stress at Month 6 calculated as the difference between the observed values and the predicted values.
resid_finstrain1	Residual Financial Strain at Wave 1	-11.36-10.38	Continuous	Float	Residual values for financial strain at Baseline calculated as the difference between the observed values and the predicted values.
resid_finstrain2	Residual Financial Strain at Wave 2	-7.85-14.19	Continuous	Float	Residual values for financial strain at Month 2 calculated as the difference between the observed values and the predicted values.

resid_finstrain3	Residual Financial Strain at Wave 3	-8.09-8.26	Continuous	Float	Residual values for financial strain at Month 3 calculated as the difference between the observed values and the predicted values.
resid_finstrain4	Residual Financial Strain at Wave 4	-9.43-12.91	Continuous	Float	Residual values for financial strain at Month 4 calculated as the difference between the observed values and the predicted values.
resid_finstrain5	Residual Financial Strain at Wave 5	-9.88-10.19	Continuous	Float	Residual values for financial strain at Month 5 calculated as the difference between the observed values and the predicted values.
resid_finstrain6	Residual Financial Strain at Wave 6	-12.81-13.04	Continuous	Float	Residual values for financial strain at Month 6 calculated as the difference between the observed values and the predicted values.
resid_socsupp1	Residual Social Support at Wave 1	-2.33-0.88	Continuous	Float	Residual values for social support at Baseline calculated as the difference between the observed values and the predicted values.
resid_socsupp2	Residual Social Support at Wave 2	-2.6-1.5	Continuous	Float	Residual values for social support at Month 2 calculated as the difference between the observed values and the predicted values.
resid_socsupp3	Residual Social Support at Wave 3	-1.45-1.62	Continuous	Float	Residual values for social support at Month 3 calculated as the difference between the observed values and the

					predicted values.
resid_socsupp4	Residual Social Support at Wave 4	-1.76-1.19	Continuous	Float	Residual values for social support at Month 4 calculated as the difference between the observed values and the predicted values.
resid_socsupp5	Residual Social Support at Wave 5	-0.99-1.3	Continuous	Float	Residual values for social support at Month 5 calculated as the difference between the observed values and the predicted values.
resid_socsupp6	Residual Social Support at Wave 6	-1.74-1.3	Continuous	Float	Residual values for social support at Month 6 calculated as the difference between the observed values and the predicted values.
r2_dep	Goodness of Fit Statistic: Depression	0.93-1653.24	Continuous	Float	Individual R^2 calculated as an indicator of how well the depression multiple group latent growth curve models fit each individual.
r2_stress	Goodness of Fit Statistic: Perceived Stress	0.63-377.01	Continuous	Float	Individual R^2 calculated as an indicator of how well the perceived stress multiple group latent growth curve models fit each individual.
r2_finstrain	Goodness of Fit Statistic: Financial Strain	1.28-2145.11	Continuous	Float	Individual R^2 calculated as an indicator of how well the financial strain multiple group latent growth curve models fit each individual.
r2_socsupp	Goodness of Fit Statistic:	0.05-51.31	Continuous	Float	Individual R^2 calculated as an

	Social Support				indicator of how well the social support multiple group latent growth curve models fit each individual.
Patterns of Missingness					
covid19	Participant completed at least one survey during COVID-19	2	Binary Categorical	Dummy	Binary indicator of completion of survey during COVID-19.
deceased_miss	Missingness Pattern: Participant died during study	2	Binary Categorical	Dummy	Binary indicator of “deceased” pattern of missingness.
dropped_miss	Missingness Pattern: Dropped from Study	2	Binary Categorical	Dummy	Binary indicator of “dropped” pattern of missingness.
itemlevel_miss	Missingness Pattern: Item-level	2	Binary Categorical	Dummy	Binary indicator of “item-level” pattern of missingness.
l2fu_miss	Missingness Pattern: Lost to Follow-up	2	Binary Categorical	Dummy	Binary indicator of “lost-to-follow-up” pattern of missingness.
onewave_miss	Missingness Pattern: Completed One Wave Only	2	Binary Categorical	Dummy	Binary indicator of “one wave only” pattern of missingness.
part_miss	Missingness Pattern: 100% Complete Participation	2	Binary Categorical	Dummy	Binary indicator of “100% completion.”
sporadic_miss	Missingness Pattern: Sporadic	2	Binary Categorical	Dummy	Binary indicator of “sporadic” pattern of missingness.

Note. Predicted values as well as residuals were computed within the Latent Variable Modeling framework using "lavPredict" (Lavaan version 0.6-9; Rosseel, 2012) following guidelines provided by Hallgren et al. (2019). In a growth model, "lavPredict" computes predicted (i.e., fitted) values for each individual based on the latent factors (i.e., latent intercept, latent slope) in the model and the observed values from the master data (Rosseel, 2012). The individual goodness of fit (GOF) statistics were calculated as an indicator of how well the multiple group latent growth curve model fit each individual as:

$$GOF_i = (y_{1i} - \hat{y}_{1i})^2 + (y_{2i} - \hat{y}_{2i})^2 + (y_{3i} - \hat{y}_{3i})^2 \ \dots + (y_{6i} - \hat{y}_{6i})^2 \qquad (1)$$

where y_{ki} is the observed recovery capital variable at a given time point k for a given individual i, and $\hat{y}_{ki}$ is the predicted value derived from the multiple group latent growth model at a given time point k for a given individual i.

Appendix D

FEATURES USED IN GENDER-SPECIFIC RANDOM FOREST MODELS

Sociodemographics Variables

Variable Name in Model	Variable	# of categories/ range	Variable Type	Encoding	Description
age	Age (numeric)	19 - 64	Continuous	Float	Age at Baseline
children	Parental Status (Has Children)	2	Binary Categorical	Dummy	Binary indicator of parental status (1=Has children)
currparole	Currently on Parole	2	Binary Categorical	Dummy	Participant was on parole at Baseline.
employed	Employed (Full-time/Part-time)	2	Binary Categorical	Dummy	Binary indicator of Employment.
hsorless	Lower Education	2	Binary Categorical	Dummy	Binary indicator of high school degree or less.
lgbtq	LGBTQ status	2	Binary Categorical	Dummy	Binary indicator of LGBTQ status.
single	Relationship Status: Single	2	Binary Categorical	Dummy	Single vs. Other Relationship Status Categories (i.e., Dating,

					Married, In a relationship, Separated, Divorced, Widowed).
white	White, non-Hispanic	2	Binary Categorical	Dummy	White, non-Hispanic vs. all other race/ethnicity categories.
Mental Health Variables					
anxiety	Anxiety Disorder Diagnosis	2	Binary Categorical	Dummy	Binary indicator of Anxiety Disorder Diagnosis.
bipolar1	Bipolar I Diagnosis	2	Binary Categorical	Dummy	Binary indicator of Bipolar I Diagnosis.
depression	Major Depression Disorder Diagnosis	2	Binary Categorical	Dummy	Binary indicator of Major Depression Disorder Diagnosis.
mhd3plus	3+ mental health disorder diagnoses	2	Binary Categorical	Dummy	Binary indicator of Major Depression Disorder Diagnosis.
ptsd	Posttraumatic Stress Disorder (PTSD) Diagnosis	2	Binary Categorical	Dummy	Binary indicator of PTSD Diagnosis.
Substance Use Variables					
daysih_max	Average Length of Stay in Home	1-1,260	Continuous	Float	Average length of stay in home (days).
lifetimerelapse	Participant experienced a relapse during lifetime	2	Binary Categorical	Dummy	Binary indicator of whether or not a participant experienced a relapse (i.e., recurrence of SUD symptoms

					after period of abstinence) at some point in their lifetime.
movedout	Participant moved out of home at some point during study	2	Binary Categorical	Dummy	Binary indicator of participant moving out of sober living home at some point during the study.
odfreq	Number of Lifetime Overdoses at Baseline	0-15	Continuous	Float	Number of overdoses experienced by participant across their lifetime, assessed at baseline.
primsub_1.0	Primary Substance of Use: Alcohol	2	Binary Categorical	Dummy	Binary indicator of Alcohol as participant's primary substance of use.
primsub_3.0	Primary Substance of Use: Cocaine	2	Binary Categorical	Dummy	Binary indicator of Cocaine as participant's primary substance of use.
recovdays90plus	Primary Substance of Use: Opioids	2	Binary Categorical	Dummy	Binary indicator of Opioids as participant's primary substance of use.
severecat	Addiction Severity	3	Categorical	Integer	Hierarchical categorical indicator of addiction severity: 1=mild (0-20) 2=moderate (21-25) 3=severe (26-30)
slh_org	Recovery Housing Organization	2	Binary Categorical	Dummy	Binary indicator of Recovery Housing Organization, where 1 = Organization 1 and 0 =

					Organization 2.
Recovery Capital Features					
pred_dep1	Predicted Depression at Wave 1	3.59-18.97	Continuous	Float	Individual predicted values for depression at Baseline.
pred_dep2	Predicted Depression at Wave 2	2.75-19.63	Continuous	Float	Individual predicted values for depression at Month 2.
pred_dep3	Predicted Depression at Wave 3	1.89-20.29	Continuous	Float	Individual predicted values for depression at Month 3.
pred_dep4	Predicted Depression at Wave 4	1.04-20.95	Continuous	Float	Individual predicted values for depression at Month 4.
pred_dep5	Predicted Depression at Wave 5	0.18-21.61	Continuous	Float	Individual predicted values for depression at Month 5.
pred_dep6	Predicted Depression at Wave 6	-0.68-22.27	Continuous	Float	Individual predicted values for depression at Month 6.
pred_stress1	Predicted Perceived Stress at Wave 1	2.05-10.87	Continuous	Float	Individual predicted values for perceived stress at Baseline.
pred_stress2	Predicted Perceived Stress at Wave 2	1.61-10.64	Continuous	Float	Individual predicted values for perceived stress at Month 2.
pred_stress3	Predicted Perceived Stress at Wave 3	1.16-10.41	Continuous	Float	Individual predicted values for perceived stress at Month 3.
pred_stress4	Predicted Perceived Stress at Wave 4	0.72-10.18	Continuous	Float	Individual predicted values for perceived stress at Month 4.
pred_stress5	Predicted Perceived	0.27-9.94	Continuous	Float	Individual predicted values for

	Stress at Wave 5				perceived stress at Month 5.
pred_stress6	Predicted Perceived Stress at Wave 6	-0.18-9.71	Continuous	Float	Individual predicted values for perceived stress at Month 6.
pred_finstrain1	Predicted Financial Strain at Wave 1	5.11-24.89	Continuous	Float	Individual predicted values for financial strain at Baseline.
pred_finstrain2	Predicted Financial Strain at Wave 2	4.87-24.79	Continuous	Float	Individual predicted values for financial strain at Month 2.
pred_finstrain3	Predicted Financial Strain at Wave 3	4.63-25.29	Continuous	Float	Individual predicted values for financial strain at Month 3.
pred_finstrain4	Predicted Financial Strain at Wave 4	4.39-25.79	Continuous	Float	Individual predicted values for financial strain at Month 4.
pred_finstrain5	Predicted Financial Strain at Wave 5	4.15-26.29	Continuous	Float	Individual predicted values for financial strain at Month 5.
pred_finstrain6	Predicted Financial Strain at Wave 6	3.91-26.79	Continuous	Float	Individual predicted values for financial strain at Month 6.
pred_socsupp1	Predicted Social Support at Wave 1	0.78-4.06	Continuous	Float	Individual predicted values for social support at Baseline.
pred_socsupp2	Predicted Social Support at Wave 2	1.17-3.8	Continuous	Float	Individual predicted values for social support at Month 2.
pred_socsupp3	Predicted Social Support at Wave 3	1.31-3.84	Continuous	Float	Individual predicted values for social support at Month 3.
pred_socsupp4	Predicted Social Support at Wave 4	1.26-3.91	Continuous	Float	Individual predicted values for social support at Month 4.
pred_socsupp5	Predicted Social Support	0.77-3.99	Continuous	Float	Individual predicted values for

	at Wave 5				social support at Month 5.
pred_socsupp6	Predicted Social Support at Wave 6	0.14-4.2	Continuous	Float	Individual predicted values for social support at Month 6.
resid_dep1	Residual Depression at Wave 1	-9.56-10.55	Continuous	Float	Residual values for Depression at Baseline calculated as the difference between the observed values and the predicted values.
resid_dep2	Residual Depression at Wave 2	-12.24-18.5	Continuous	Float	Residual values for Depression at Month 2 calculated as the difference between the observed values and the predicted values.
resid_dep3	Residual Depression at Wave 3	-11.26-12.25	Continuous	Float	Residual values for Depression at Month 3 calculated as the difference between the observed values and the predicted values.
resid_dep4	Residual Depression at Wave 4	-11.61-12.71	Continuous	Float	Residual values for Depression at Month 4 calculated as the difference between the observed values and the predicted values.
resid_dep5	Residual Depression at Wave 5	-8.52-11.99	Continuous	Float	Residual values for Depression at Month 5 calculated as the difference between the observed values and the predicted values.
resid_dep6	Residual Depression at Wave 6	-15.53-9.78	Continuous	Float	Residual values for Depression at Month 6 calculated as the difference between the observed

					values and the predicted values.
resid_stress1	Residual Perceived Stress at Wave 1	-4.37-5.47	Continuous	Float	Residual values for perceived stress at Baseline calculated as the difference between the observed values and the predicted values.
resid_stress2	Residual Perceived Stress at Wave 2	-4.77-6.61	Continuous	Float	Residual values for perceived stress at Month 2 calculated as the difference between the observed values and the predicted values.
resid_stress3	Residual Perceived Stress at Wave 3	-3.88-5.44	Continuous	Float	Residual values for perceived stress at Month 3 calculated as the difference between the observed values and the predicted values.
resid_stress4	Residual Perceived Stress at Wave 4	-5.6-6.45	Continuous	Float	Residual values for perceived stress at Month 4 calculated as the difference between the observed values and the predicted values.
resid_stress5	Residual Perceived Stress at Wave 5	-3.39-6.96	Continuous	Float	Residual values for perceived stress at Month 5 calculated as the difference between the observed values and the predicted values.
resid_stress6	Residual Perceived Stress at Wave 6	-3.66-4.04	Continuous	Float	Residual values for perceived stress at Month 6 calculated as the difference between the observed values and the predicted values.
resid_finstrain1	Residual Financial	-11.36-10.38	Continuous	Float	Residual values for financial strain

	Strain at Wave 1				at Baseline calculated as the difference between the observed values and the predicted values.
resid_finstrain2	Residual Financial Strain at Wave 2	-7.85-14.19	Continuous	Float	Residual values for financial strain at Month 2 calculated as the difference between the observed values and the predicted values.
resid_finstrain3	Residual Financial Strain at Wave 3	-8.09-8.26	Continuous	Float	Residual values for financial strain at Month 3 calculated as the difference between the observed values and the predicted values.
resid_finstrain4	Residual Financial Strain at Wave 4	-9.43-12.91	Continuous	Float	Residual values for financial strain at Month 4 calculated as the difference between the observed values and the predicted values.
resid_finstrain5	Residual Financial Strain at Wave 5	-9.88-10.19	Continuous	Float	Residual values for financial strain at Month 5 calculated as the difference between the observed values and the predicted values.
resid_finstrain6	Residual Financial Strain at Wave 6	-12.81-13.04	Continuous	Float	Residual values for financial strain at Month 6 calculated as the difference between the observed values and the predicted values.
resid_socsupp1	Residual Social Support at Wave 1	-2.33-0.88	Continuous	Float	Residual values for social support at Baseline calculated as the difference between the observed values and the predicted values.

resid_socsupp2	Residual Social Support at Wave 2	-2.6-1.5	Continuous	Float	Residual values for social support at Month 2 calculated as the difference between the observed values and the predicted values.
resid_socsupp3	Residual Social Support at Wave 3	-1.45-1.62	Continuous	Float	Residual values for social support at Month 3 calculated as the difference between the observed values and the predicted values.
resid_socsupp4	Residual Social Support at Wave 4	-1.76-1.19	Continuous	Float	Residual values for social support at Month 4 calculated as the difference between the observed values and the predicted values.
resid_socsupp5	Residual Social Support at Wave 5	-0.99-1.3	Continuous	Float	Residual values for social support at Month 5 calculated as the difference between the observed values and the predicted values.
resid_socsupp6	Residual Social Support at Wave 6	-1.74-1.3	Continuous	Float	Residual values for social support at Month 6 calculated as the difference between the observed values and the predicted values.
r2_dep	Goodness of Fit Statistic: Depression	0.93-1653.24	Continuous	Float	Individual R^2 calculated as an indicator of how well the depression multiple group latent growth curve models fit each individual.
r2_stress	Goodness of Fit Statistic: Perceived	0.63-377.01	Continuous	Float	Individual R^2 calculated as an indicator of how well the perceived

	Stress				stress multiple group latent growth curve models fit each individual.
r2_finstrain	Goodness of Fit Statistic: Financial Strain	1.28-2145.11	Continuous	Float	Individual R^2 calculated as an indicator of how well the financial strain multiple group latent growth curve models fit each individual.
r2_socsupp	Goodness of Fit Statistic: Social Support	0.05-51.31	Continuous	Float	Individual R^2 calculated as an indicator of how well the social support multiple group latent growth curve models fit each individual.
Patterns of Missingness					
covid19	Participant completed at least one survey during COVID-19	2	Binary Categorical	Dummy	Binary indicator of completion of survey during COVID-19.
deceased_miss	Missingness Pattern: Participant died during study	2	Binary Categorical	Dummy	Binary indicator of "deceased" pattern of missingness.
dropped_miss	Missingness Pattern: Dropped from Study	2	Binary Categorical	Dummy	Binary indicator of "dropped" pattern of missingness.
itemlevel_miss	Missingness Pattern: Item-level	2	Binary Categorical	Dummy	Binary indicator of "item-level" pattern of missingness.
l2fu_miss	Missingness Pattern: Lost to Follow-up	2	Binary Categorical	Dummy	Binary indicator of "lost-to-follow-up" pattern of missingness.

onewave_miss	Missingness Pattern: Completed One Wave Only	2	Binary Categorical	Dummy	Binary indicator of “one wave only” pattern of missingness.
part_miss	Missingness Pattern: 100% Complete Participation	2	Binary Categorical	Dummy	Binary indicator of “100% completion.”
sporadic_miss	Missingness Pattern: Sporadic	2	Binary Categorical	Dummy	Binary indicator of “sporadic” pattern of missingness.

Note. Predicted values as well as residuals were computed within the Latent Variable Modeling framework using “lavPredict” (Lavaan version 0.6-9; Rosseel, 2012) following guidelines provided by Hallgren et al. (2019). In a growth model, “lavPredict” computes predicted (i.e., fitted) values for each individual based on the latent factors (i.e., latent intercept, latent slope) in the model and the observed values from the master data (Rosseel, 2012). The individual goodness of fit (GOF) statistics were calculated as an indicator of how well the multiple group latent growth curve model fit each individual as:

$$GOF_i = (y_{1i} - \hat{y}_{1i})^2 + (y_{2i} - \hat{y}_{2i})^2 + (y_{3i} - \hat{y}_{3i})^2 \ldots + (y_{6i} - \hat{y}_{6i})^2 \quad (1)$$

where y_{ki} is the observed recovery capital variable at a given time point k for a given individual i, and $\hat{y}_{ki}$ is the predicted value derived from the multiple group latent growth model at a given time point k for a given individual i.