Impact of MGM Springfield on Gambling Attitudes, Participation and Problem Gambling

November 13, 2020

SEIGMA
SOCIAL AND ECONOMIC IMPACTS OF GAMBLING IN MASSACHUSETTS

UNIVERSITY OF MASSACHUSETTS SCHOOL OF PUBLIC HEALTH AND HEALTH SCIENCES

Authorship and Acknowledgements

Authorship

Rachel A. Volberg, Research Professor at the School of Public Health and Health Sciences, University of Massachusetts Amherst and Principal Investigator on the SEIGMA project. Dr. Volberg is the lead author of this report.

Martha Zorn, SEIGMA Data Manager, School of Public Health and Health Sciences, University of Massachusetts at Amherst. Ms. Zorn was responsible for data cleaning, data management, and data analysis and contributed to all sections of the report.

Valerie Evans, SEIGMA Project Manager and Biostatistician. Ms. Evans assisted in drafting the report and provided technical assistance and quality assurance.

Edward J. Stanek, Professor Emeritus, University of Massachusetts Amherst School of Public Health and Health Sciences, contributed sections of the report related to weighting.

Robert J. Williams, Professor in the Faculty of Health Sciences at the University of Lethbridge in Alberta, Canada, and Co-Principal Investigator on the SEIGMA project. Dr. Williams contributed to all sections of the report.

Acknowledgements

Initial financial support for Social and Economic Impacts of Gambling in Massachusetts (SEIGMA) study came in 2013 from the Massachusetts Gaming Commission under ISA MGC10500003UMS15A. The multi-year project was competitively bid via the Massachusetts Gaming Commission Request for Response (MGC-RA-2012) for Research Services and awarded to the University of Massachusetts Amherst in April 2013. In June 2019 the Massachusetts Gaming Commission issued a subsequent Request for Response (BD-19-1068-1700-1-40973) for Research Services and the University of Massachusetts Amherst was awarded the contract effective January 2020.

The population surveys on which the analyses in this report rest could not have been completed without the cooperation and good will of the thousands of Massachusetts residents who agreed to participate. We are grateful to the many individuals at NORC at the University of Chicago who helped in collecting the data for the surveys.

We would like to thank the members of the Massachusetts Gaming Commission's Research Review Committee (RRC). Members of this committee represent a range of perspectives and their careful review of draft versions of this report contributed to its clarity as well as utility to multiple audiences.

Finally, we would like to thank the staff and members of the Massachusetts Gaming Commission, for their thoughtful input and clear guidance over the course of the SEIGMA project. The Commission's broad vision for the expansion of gambling in Massachusetts and commitment to the research needed to maximize the benefits and minimize harms related to gambling in the Commonwealth made this project possible.

SUGGESTED CITATION:

Volberg, R.A., Zorn, M., Evans, V., Stanek, E.J., Williams, R.J. (2020). *Impact of MGM Springfield on Gambling Attitudes, Participation, and Problem Gambling*. Amherst, MA: School of Public Health and Health Sciences, University of Massachusetts Amherst**.**

A PDF OF THIS REPORT CAN BE DOWNLOADED AT: **www.umass.edu/seigma**

Table of Contents

Authorship and Acknowledgements i
List of Figures iii
List of Tables iii
Executive Summary iv
Acronyms and Abbreviations viii
Introduction 1
Overview of Methods 3
Results 11
Discussion 29
References 33
Appendix A: Weighting Calculations 37
Appendix B: Tabular Results 96

List of Figures

Figure 1. Change in Attitudes about Gambling from Baseline (2015) to Follow-up (2019) 14
Figure 2. Changes in Gambling Participation from Baseline (2015) to Follow-up (2019) 17
Figure 3. Changes in Gambling Participation by Education... 19
Figure 4. Changes in Gambling Participation by Income.. 21
Figure 5. Change in Problem Gambling Prevalence 24
Figure 6. Changes in PPGM by Income.. 25
Figure 7. Change in Awareness of Problem Gambling Programs.. 27

List of Tables

Table 1. Demographics of Springfield Baseline and Follow-up Target Survey Sample 7
Table 2. Attitudes: Targeted Springfield Survey (Weighted).. 97
Table 3. Gambling Behaviors: Targeted Springfield Survey (Weighted) 98
Table 4. Gambling Behaviors: Targeted Springfield Survey (Weighted) by Gender... 99
Table 5. Gambling Behaviors: Targeted Springfield Survey (Weighted) by Age 100
Table 6. Gambling Behaviors: Targeted Springfield Survey (Weighted) by Race/Ethnicity 101
Table 7. Gambling Behaviors: Targeted Springfield Survey (Weighted) by Education 102
Table 8. Gambling Behaviors: Targeted Springfield Survey (Weighted) by Income... 103
Table 9. Gambling Intensity: Targeted Springfield Survey (Weighted) 104
Table 10. PPGM (Collapsed): Targeted Springfield Survey (Weighted)... 105
Table 11. PPGM: Targeted Springfield Survey (Weighted) by Gender.. 105
Table 12. PPGM: Targeted Springfield Survey (Weighted) by Age.. 106
Table 13. PPGM: Targeted Springfield Survey (Weighted) by Race/Ethnicity 107
Table 14. PPGM: Targeted Springfield Survey (Weighted) by Education... 108
Table 15. PPGM: Targeted Springfield Survey (Weighted) by Income.. 109
Table 16. Awareness of Media Campaigns and Programs: Targeted Springfield Survey (Weighted).................. 110
Table 17. Treatment: Targeted Springfield Survey (Weighted) 110

Executive Summary

Background

The Social and Economic Impacts of Gambling in Massachusetts (SEIGMA) study is a comprehensive, multi-year investigation of the impacts of introducing casino gambling in the Commonwealth. Beginning in 2013, the SEIGMA research team has collected extensive baseline and follow-up data on the social and economic changes related to the introduction of casino gambling in Massachusetts. The study established baselines for all social and economic variables that may be affected by expanded gaming, and the team now collects, analyzes, and reports each year to identify the actual impacts in the casino host and surrounding communities, providing key information to policymakers and other interested stakeholders.

This report focuses on one aspect of the broader SEIGMA study and summarizes findings from baseline and follow-up targeted population surveys carried out in Springfield and surrounding communities. The baseline targeted survey was conducted in 2015, soon after the announcement of the award of license to MGM Resorts International. The follow-up targeted survey was conducted in 2019, one year after the opening of MGM Springfield in June 2018. The Springfield Baseline and Follow-up Targeted Population Surveys largely used the same methodology as the state-wide Baseline General Population Survey (BGPS) that was conducted in 2013 and 2014. Findings from these surveys are important to help understand changes in attitudes toward gambling, gambling behavior, the prevalence of problem gambling, and awareness of and involvement in problem gambling services in the wake of the introduction of a major gambling venue. The findings will contribute to the development of strategies to minimize gambling-related harm and bring the greatest possible benefit of expanded gaming to the people of the Commonwealth.

Methods

The SEIGMA team obtained a probability sample of households in Springfield and surrounding communities and allowed survey respondents to complete the survey online, on paper, or by telephone. The Baseline Targeted Population Survey in Springfield (BTPS-S) took place between February 2015 and June 2015, had a response rate of 31.7%, and included a final sample of 1,131 Western Massachusetts residents aged 18 and over. The majority of questionnaires were self-administered (90%) and 2.5% of the questionnaires were completed in Spanish. The Follow-up Targeted Population Survey in Springfield (FTPS-S) took place between October 2019 and January 2020, had a response rate of 16.7%, and included a final sample of 1,134 Western Massachusetts residents aged 18 and over. The majority of questionnaires were self-administered (89%) and 14.7% of the questionnaires were completed in Spanish. The data from both surveys was weighted to align the sample more closely with the target population. In addition to the surveys, the SEIGMA team conducted interviews with key informants in Springfield to gain an on-the-ground understanding from local experts about conditions in Springfield prior to hosting a casino and after the casino opened.

It is important to emphasize that the targeted surveys in Springfield and surrounding communities were cross-sectional 'snapshots' of gambling and problem gambling at single points in time. This limits our ability to draw any cause-and-effect conclusions from associations reported between gambling participation, gambling problems, and other variables in these surveys.

Key Findings[1]

Attitudes toward Gambling

There were several significant changes in attitudes toward gambling among residents of Springfield and surrounding communities between 2015 and 2019. First, compared to 2015, more residents surveyed in 2019 believed that the availability of gambling in Massachusetts was too high. Second, the majority of residents in both 2015 and 2019 believed that the harm of gambling to society outweighed the benefits with a significantly higher proportion feeling this way in 2019. Third, compared to 2015, more residents viewed the importance of gambling as a recreational activity as "not at all important." Finally, compared to 2015, fewer residents surveyed in 2019 viewed the impact of expanded gambling in Massachusetts as harmful and more residents held a neutral view. Taken together, these changes suggest that overall perceptions of gambling among residents of Springfield and surrounding communities became somewhat more negative but also less polarized between 2015 and 2019.

Gambling Participation

Between 2015 and 2019, overall gambling participation changed very little. There was a statistically significant increase in past year participation in daily lottery games that was likely due to changes in the question wording. Past year gambling at out-of-state casinos did not change significantly but there was a statistically significant increase in any casino gambling in the past year. This change was driven by the greater proportion of residents of Springfield and surrounding communities who gambled at both Massachusetts and out-of-state casinos in 2019. Beyond past year daily lottery play and overall casino gambling, there were no statistically significant changes in gambling behavior among residents of Springfield and surrounding communities between 2015 and 2019.

Given the lack of changes in past year participation in most specific forms of gambling, we felt it was important to explore whether there were changes in overall gambling participation, overall lottery participation, and overall casino gambling in Massachusetts and out-of-state by specific demographic groups. There were no significant changes in overall gambling or overall lottery participation by **gender**, **race/ethnicity** and among adults aged 50 and over. The rate of overall casino gambling was significantly higher in 2019 compared to 2015 among males and females, among Blacks/Hispanics/Asians and Whites/Other, and among individuals aged 50 to 64 and those aged 65 and over. In contrast to most other demographic groups, the rate of overall casino gambling did not change significantly between 2015 and 2019 among adults aged 18 to 34 and those aged 35 to 49. Among adults aged 18 to 34, the rate of overall lottery participation was significantly lower in 2019 compared to 2015 (45.8% in2019 compared to 62.5% in 2015).

When it comes to **educational attainment**, rates of overall gambling and overall lottery participation changed very little among residents of Springfield and surrounding communities with different levels of education. The rate of overall casino gambling was significantly higher among those with a high school education or less and among those who attended or graduated from college in 2019 compared to 2015. This was not the case among those with graduate level education. With regard to **income**, there were no statistically significant changes in rates of overall gambling, overall lottery participation, or overall casino gambling among residents of Springfield and surrounding communities with annual household incomes of $50,000 or higher. Among residents with annual household incomes lower than $50,000, rates of overall gambling and overall casino gambling were significantly higher in 2019 compared to 2015.

[1] Only differences that are statistically significant at the 0.01 level are included in the Executive Summary and in the body of the report.

Problem Gambling

One of the main negative social impacts of expanded gambling availability tends to be an increase in problem gambling. In epidemiological research, prevalence is a measure of the number of individuals in the population with a disorder at one point in time. In problem gambling prevalence surveys, individuals are classified on the basis of their responses to a valid and reliable problem gambling instrument. The Problem and Pathological Gambling Measure (PPGM) serves as the primary instrument to assess problem gambling in the SEIGMA study. Based on the PPGM, there was no change in the prevalence of at-risk and problem gambling between 2015 and 2019 among residents of Springfield and surrounding communities.

We estimate that between 42,074 (9.3%) and 70,123 (15.5%) residents of Springfield and surrounding communities were at-risk for or experiencing a gambling problem in 2015. In 2019, we estimate that between 36,421 (8.0%) and 63,281 (13.9%) adult residents of Springfield and surrounding communities were at-risk for or experiencing a gambling problem.

Problem gambling prevalence rates can vary significantly across important subgroups in the population. We examined differences and changes in problem gambling prevalence across the same five demographic groups discussed earlier in relation to gambling participation. The only statistically significant change was a decrease in non-gambling (accompanied by a substantial but not significantly higher rate of recreational gambling) in 2019 among residents of Springfield and surrounding communities with annual household incomes under $50,000.

Awareness and Utilization of Problem Gambling Programs

Previous research has found that many people experiencing gambling problems recover without the aid of professional treatment. Indeed, the literature indicates that the number of people who have recovered on their own may greatly exceed the number of people who ever seek treatment. These findings highlight the importance of increasing public awareness and encouraging changes in attitudes and behavior among individuals experiencing mild or moderate difficulties to reduce their progression toward more severe gambling-related problems.

Between 2015 and 2019, there was a statistically significant reduction in the proportion of residents of Springfield and surrounding communities who indicated that they were aware of media campaigns to prevent problem gambling in Massachusetts in the past year. Almost half of the respondents in the 2015 survey (47.9%) were aware of problem gambling prevention media campaigns in 2015 but the proportion dropped to 32.1% in 2019. Awareness of problem gambling prevention programs other than media campaigns was lower than awareness of media campaigns in both surveys and the change between 2015 and 2019 was not statistically significant. It is possible that changes in the administration of problem gambling services in Massachusetts beginning in 2016, along with the end of heated public discussion of the casino issue in Western Massachusetts, contributed to these changes.

Responses to email and telephone queries to two of the three Gamblers Anonymous meetings in the Springfield area indicated that these meetings have taken place for many years with little change in the number of attendees following the opening of the casino. There has been a much greater impact on these meetings from COVID-19, which has limited the number of participants to 10 rather than the more usual 25-30 attendees.

Discussion and Future Directions

The relationship between proximity, and thereby availability, of gambling venues and the likelihood of experiencing gambling problems has long been debated. The availability, or exposure, theory suggests that an increase in the availability of gambling venues to a population would lead to an increase in gambling-related harms, particularly gambling problems. Alternatively, the adaptation theory suggests that there will be an initial increase in gambling-related problems upon exposure of a resident population to a new gambling venue with the effects abating over time.

Replication surveys that examine changes in problem gambling prevalence in the same jurisdiction over time provide a direct test of exposure versus adaptation. Our findings from the replication survey carried out in Springfield and surrounding communities (like our findings from the Plainville region) suggest that the Massachusetts population is far from naïve when it comes to casino gambling. States surrounding Massachusetts, including Rhode Island, Connecticut, and New York, have had casino gambling for decades prior to the introduction of casino gambling in Massachusetts. Following this initial exposure, any effects may have abated over time, even in a population that has experienced recent local gambling expansions. In our view, the findings from this study suggest that population adaptation has already occurred as no increased risk of harms associated with casino gambling was identified.

There are additional factors that may have contributed to this perceived adaptation. An increase in public awareness through media or public health campaigns, at least during the period before the casino opened, may have raised awareness of the potential harms and, subsequently, may have reduced involvement by at-risk individuals. The expansion of treatment services for those individuals who do experience gambling problems may have contributed to increased rates of recovery and fewer relapses. Regulatory or industry measures instituted to curtail gambling harms and increase consumer safety, such as casino self-exclusion programs or the Massachusetts Gaming Commission's GameSense program, may also have prevented some at-risk individuals from developing gambling problems.

While the overall results of the surveys in Springfield and surrounding communities are reassuring, there are concerns about how specific demographic groups in the region may be affected in the future. These groups include individuals with lower educational attainment and individuals with annual household incomes under $50,000. Changes in overall gambling participation and overall casino participation within these groups suggest that individuals with lower education and lower income may be vulnerable to experiencing gambling harms or developing gambling problems because the location of MGM Springfield has made it easier for them to engage in a type of gambling with which they have had relatively little experience in the past. It will be important to direct prevention and treatment resources toward these groups going forward as well as to assess at-risk and problem gambling rates in these groups in the future.

Acronyms and Abbreviations

ABS – Address Based Sampling
BGPS – Baseline General Population Survey
BTPS – Baseline Targeted Population Survey
CASRO – Council of American Survey Research Organizations
CATI – Computer Assisted Telephone Interview
CI – Confidence Interval
CPGI – Canadian Problem Gambling Index
DSM – Diagnostic and Statistical Manual of Mental Disorders
EGM – Electronic Gaming Machine
FTPS – Follow-up Targeted Population Survey
MAGIC – Massachusetts Gambling Impact Cohort Study
MGC – Massachusetts Gaming Commission
NORC – NORC at the University of Chicago (formerly National Opinion Research Center)
PPGM – Problem and Pathological Gambling Measure
PUMS – Public Use Microdata Sample
RSE – Relative Standard Error
SAQ – Self Administered Questionnaire
SE – Standard Error
SEIGMA – Social and Economic Impacts of Gambling in Massachusetts
SOGS – South Oaks Gambling Screen
UMass – University of Massachusetts

Introduction

In November 2011, an Act Establishing Expanded Gaming in the Commonwealth was passed by the Legislature and signed by Governor Deval Patrick (Chapter 194 of the Acts of 2011). This legislation permitted casinos and slot parlors to be introduced in Massachusetts under the regulatory auspices of the Massachusetts Gaming Commission (MGC). The Expanded Gaming Act also required the MGC to establish an "annual research agenda" to understand the social and economic effects of casino gambling. In March 2013, the MGC awarded a contract to a team at the University of Massachusetts Amherst to conduct this research. This research project is known as the Social and Economic Impacts of Gambling in Massachusetts (SEIGMA) study. In December 2019, the MGC awarded a contract to the same team to continue the study for several more years.

SEIGMA was originally envisioned as a discrete before/after evaluation of the impact of the introduction of casinos into Massachusetts. However, the gradual introduction of the new casinos in Massachusetts over an extended period of time (2015 - 2019) meant that waiting until after all of the casinos opened might miss impacts arising as a result of the earlier casino introductions. By the same token, some impacts can take several years to fully manifest and so a singular evaluation in the midst of the introductory phase of casino gambling in Massachusetts would miss these changes.

The SEIGMA study encompasses most of the essential elements contained in the MGC research agenda, which includes:

- Understanding the social and economic effects of expanded gambling;
- Implementing a baseline study of problem gambling and the existing prevention and treatment programs that address its harmful consequences; and
- Obtaining scientific information relative to the epidemiology and etiology of gambling.

The SEIGMA study uses a collaborative orientation, a state-of-the-art analytical framework, a mixed methods research strategy, and a comprehensive approach that establishes the impacts of casino gambling at state, regional, and local levels. Importantly, the study is unique in obtaining information about gambling involvement and problem gambling prevalence prior to the introduction of casino gambling and re-assessing these behaviors soon after the casinos have opened.

This report focuses on one aspect of the broader SEIGMA study and summarizes findings from the baseline and follow-up targeted population surveys carried out in Springfield and surrounding communities. The baseline targeted survey was conducted in 2015, soon after the announcement of the award of license to MGM Resorts International. The follow-up targeted survey was conducted in 2019, one year after the opening of MGM Springfield in June 2018. Findings from these surveys are an important component in the effort to understand changes in gambling behavior and problem gambling prevalence in the wake of the introduction of a major gambling venue and to develop strategies to minimize gambling-related harm and bring the greatest possible benefit of expanded gaming to the people of the Commonwealth.

Rationale for Conducting Population Surveys of Gambling

The gambling studies field has changed considerably over the last 30 years. In the 1980s and early 1990s, when the first surveys of gambling and problem gambling were carried out, policy makers were simply interested in finding out how many people experiencing gambling problems there were in a jurisdiction in order to fund and design treatment services for individuals with gambling-related difficulties. Since that time, the goals for gambling prevalence research have become more complex.

The growth of legal, commercial gambling has been accompanied by an increase in stakeholders with interests in and concerns about the gambling industry and how it affects individuals, families, and communities. Policy makers, planners, and government agencies are concerned with a broad range of gambling behaviors in the population, as well as with the balance of positive and negative impacts that may accompany the increased availability of gambling. Regulators and operators are interested in how to manage funds appropriately to address the issue of problem gambling while still maintaining a viable commercial industry. Public health professionals, social scientists, and healthcare providers are interested in identifying ways to minimize risks to specific subgroups in the population. Other professionals, such as economists, law enforcement professionals and the banking, insurance, and credit card industries, are interested in the relationship between gambling, indebtedness, bankruptcy, and crime. There is also interest in the extent of the profits flowing to the gambling industry from gambling by people experiencing gambling problems. Treatment professionals and non-profit organizations are focused on developing appropriate treatment services and in judiciously allocating the resources that flow to the mental health and addictions field. Finally, there is growing interest in prevention strategies and interventions for minimizing gambling-related harms.

Population surveys of gambling have become an essential component in establishing and monitoring legal gambling (Volberg, 2004; Volberg & Wray, 2013; R. J. Williams, Volberg, R.A., & Stevens, R.M.G., 2012; Young, 2013). Results of these surveys can be used to shape public awareness campaigns using targeted messages to prompt changes in attitudes and behavior in vulnerable subgroups in the population. Population surveys can also inform the development of treatment services for individuals with gambling problems, through identification of the number and characteristics of individuals likely to seek help. Population surveys have the potential to improve how gambling problems are identified and how communities respond. Finally, population surveys have value in advancing understanding of the risk factors associated with gambling problems—information needed in the development of evidence-based gambling interventions, regulations, and policies.

Purpose of Report

This report presents a comprehensive compilation of descriptive statistical (univariate and bivariate) results from the baseline and follow-up targeted surveys in Springfield and surrounding communities. The report is organized into several sections for clarity of presentation. Following this *Introduction*, an *Overview of Methods* details how we conducted the survey. The next section, *Results*, presents findings from the surveys in the following areas:

- Attitudes toward gambling
- Gambling behavior
- Prevalence of problem gambling
- Relationship of particular forms of gambling to problem gambling in Western MA
- Comparing recreational, at-risk, and problem gamblers
- Attitudes toward, awareness of, and involvement in problem gambling services

The report concludes with a summary of the findings of the study. There are several appendices to the report, including a detailed explanation of the weighting of the samples, a copy of the questionnaires that were used, and comprehensive tables that summarize the findings from most of the items in the surveys.

Overview of Methods

In addition to the statewide Baseline General Population Survey (BGPS) conducted in 2013/2014, 'Targeted Population Surveys' have been conducted in the geographic areas where new casinos and the slot parlor have been built. These targeted areas include the host community where the casino is located as well as the surrounding communities which are defined as municipalities proximate to a host community and which the MGC deems likely to experience impacts from the new venue. There are both 'Baseline Targeted Population Surveys' (before the casino has opened) and 'Follow-Up Targeted Population Surveys' (originally planned after the casino had been opened for one year). For the most part, the same methodology utilized in the BGPS was employed in the Baseline and Follow-up Targeted Population Surveys (BTPS and FTPS). In this section, we provide an overview of the methods employed in the Baseline and Follow-up Targeted Population Surveys in Springfield. We also provide information about the key informant interviews that were conducted in Springfield to gain an on-the-ground understanding from local experts about conditions in Springfield prior to and after the construction of the casino.

Baseline Targeted Population Survey – Springfield

The Baseline Targeted Population Survey in Springfield (BTPS-S) was conducted by NORC at the University of Chicago under contract to the University of Massachusetts Amherst, School of Public Health and Health Sciences. The goals of the survey were to establish a baseline level of gambling participation and problem gambling prevalence and to assess awareness and utilization of problem gambling services prior to the opening of MGM Springfield. The survey protocol was reviewed and approved separately by NORC's Institutional Review Board and by the University of Massachusetts Amherst Institutional Review Board.

The BTPS-S used address-based probability sampling to ensure that all Springfield and surrounding communities' households had a known probability of selection into the sample. Within each sampled dwelling unit, the adult with the most recent birthday was selected as the survey respondent.

The BTPS-S involved a multi-modal approach to provide eligible respondents with different ways to complete the questionnaire. NORC mailed letters to all selected addresses and subsequent postcards inviting the adult (18+) household member with the most recent birthday to complete an online survey. The letter contained a $1 incentive and offered respondents a $10 Amazon gift-code if the survey was completed within 14 days. A thank-you or reminder postcard was mailed out one week after the advance letter. Two weeks later, a second letter was mailed out encouraging respondents to complete the survey online and including the web link and PIN to access the survey. If respondents had not completed the survey online five weeks after the advance letter, they were sent a paper-and-pencil questionnaire along with an explanatory letter, a $5 incentive, and a return envelope. Two weeks later, a thank-you or reminder postcard was mailed out. Two weeks later, households received a second reminder postcard thanking those who had previously completed the survey and reminding non-respondents to complete the survey. Every address that failed to complete the survey via mail or online and whose household had been matched with a landline telephone number was then called and given the opportunity to complete the survey over the telephone. Telephone interviews were conducted by trained interviewers using a CATI system.

Data collection began in February 2015 and ended in June 2015. The CASRO response rate was 31.7% and the final sample included 1,131 Western Massachusetts residents aged 18 and over. The majority of questionnaires were self-administered, with 47% completed online and 43% completed using the paper-and-pencil questionnaire. The survey was offered in English and Spanish. A total of 28 questionnaires or telephone interviews (2.5%) were completed in Spanish.

Follow-up Targeted Population Survey – Springfield

Like the baseline targeted survey, the Follow-up Targeted Population Survey in Springfield (FTPS-S) was conducted by NORC at the University of Chicago. The goals of the survey were to identify levels of gambling participation and problem gambling prevalence and to assess awareness and utilization of problem gambling services approximately one year after the opening of MGM Springfield. The survey protocol was reviewed and approved separately by NORC's Institutional Review Board and by the University of Massachusetts Amherst Institutional Review Board.

The FTPS-S used address-based probability sampling to ensure that all Springfield and surrounding communities' households had a known probability of selection into the sample. To additionally ensure adequate representation of specific demographic groups, targeted list samples were used in combination with the USPS Delivery Sequence File. Oversampling of Hispanic and African American households was based on both targeted list sampling and area-based stratification of Census tracts. Within each sampled dwelling unit, the adult with the most recent birthday was selected as the survey respondent.

The FTPS-S involved a multi-modal approach to provide eligible respondents with different ways to complete the questionnaire. NORC mailed letters to all selected addresses and subsequent postcards inviting the adult (18+) household member with the most recent birthday to complete an online survey. The letter contained a $1 incentive and offered respondents a $10 Amazon gift-code if the survey was completed within 14 days. A thank-you or reminder postcard was mailed out one week after the advance letter. Two weeks later, a second letter was mailed out encouraging respondents to complete the survey online and including the web link and PIN to access the survey. If respondents had not completed the survey online five weeks after the advance letter, they were sent a paper-and-pencil questionnaire along with an explanatory letter, a $5 incentive, and a return envelope. Two weeks later, a thank-you or reminder postcard was mailed out. Two weeks later, households received a second reminder postcard thanking those who had previously completed the survey and reminding non-respondents to complete the survey. Every address that failed to complete the survey via mail or online and whose household had been matched with a landline telephone number was then called and given the opportunity to complete the survey over the telephone. Telephone interviews were conducted by trained interviewers using a CATI system.

Data collection began in October 2019 and ended in January 2020. When the initial fielded sample resulted in fewer completed interviews than projected, the SEIGMA research team and NORC agreed to field additional sample to achieve the targeted sample size while recognizing that this would likely negatively affect the overall response rate for the survey. Additional sample for the web survey modality only was fielded in December 2019. The final CASRO response rate was 16.7% and the final sample included 1,134 Western Massachusetts residents aged 18 and over. The majority of questionnaires were self-administered, with 72% completed online and 27% completed using the paper-and-pencil questionnaire. A total of 167 questionnaires or telephone interviews (14.7%) were completed in Spanish.[2]

[2] The higher proportion of people who completed the survey in Spanish in 2019 (compared with 2015) was almost certainly due to the use of targeted lists. Targeted lists were used to ensure representativeness of important subgroups in the population. This approach was successful but occurred at the expense of achieving a higher response rate. Corrective weighting was used to align the demographics of the achieved samples in both surveys with the known characteristics of the population.

Questionnaire

The questionnaire for both the BTPS-S and the FTPS-S included sections on recreation, physical, and mental health behaviors, alcohol and drug use, attitudes toward gambling, gambling participation, gambling motivations, awareness of problem gambling services, gambling-related problems, and demographics. Gambling participation was assessed by asking about past year frequency of participation in 11 different types of gambling:

- large jackpot lottery tickets
- instant tickets or pull tabs
- daily lottery games (in 2015, this included Keno and Jackpot Poker; in 2019, this included Mass Cash, All or Nothing, and Numbers Game)
- raffle tickets
- betting money on sporting events (this includes sports pools)
- bingo
- casino, racino, or slots parlor outside of Massachusetts
- horse racing (on-site track or an off-track site)
- betting money against other people on things such as card games, golf, pool, darts, bowling, video games, board games, or poker outside of a casino
- high risk stocks, options, or futures or day trade on the stock market
- gambling online (including playing poker, buying lottery tickets, betting on sports, bingo, slots or casino table games for money, or playing interactive games for money)

Five questions were added to the FTPS-S questionnaire to assess participation in specific casino gambling activities in greater detail compared to the BTPS-S and the BGPS. Respondents were asked about gambling at a casino or slot parlor in Massachusetts, frequency of gambling at a casino or slot parlor in Massachusetts, money spent at Massachusetts casinos or slot parlors in a typical month, spending money on electronic gambling machines either at a casino or online, and betting money on casino table games either at a casino or online.

All participants who reported gambling once a month or more on one or more types of gambling were administered the Problem and Pathological Gambling Measure (PPGM) (R. J. Williams & Volberg, 2010, 2014). The PPGM has higher sensitivity, specificity, and classification accuracy for population assessment of problem gambling compared to older instruments such as the Problem Gambling Severity Index (PGSI), Diagnostic and Statistical Manual of Mental Disorders—Fourth Edition (DSM-IV), and the South Oaks Gambling Screen (SOGS).

Based on responses to the PPGM, a person was categorized as a Non-Gambler if he or she reported no past year participation in any form of gambling (with the exception of high-risk stocks). A person was categorized as a Recreational Gambler if he or she reported participating in one or more types of gambling in the past year but no problem gambling symptomatology and frequency of gambling and gambling expenditure were below levels reported by Problem and Pathological Gamblers. A person was categorized as an At-Risk Gambler if he or she reported participating in one or more types of gambling in the past year *and* reported one or more symptoms of problem gambling. Alternatively, a person could be classified as an At-Risk Gambler if their frequency of gambling and gambling losses were equal to or greater than the median reported for Problem and Pathological Gamblers. A person was categorized as a Problem Gambler if he or she reported: gambling at least once a month on one or more types of gambling; a Problems Score of 1 or higher; an Impaired Control Score of 1 or higher; and a Total Score of 2 to 4. Alternatively, a person could receive this designation if they had a Total Score of 3 or higher plus a frequency of gambling and reported gambling loss that was equal to or greater than the median for Problem and Pathological Gamblers. A person was categorized as a Pathological Gambler if he or she reported: gambling at least once a month on one or more types of gambling; a Problems Score of 1 or higher; an

Impaired Control Score of 1 or higher; and a Total Score of 5 or higher. In the statistical analyses, At-Risk, Problem and Pathological Gamblers were collapsed into one group due to small cell sizes.

Weighting and Imputation

The ultimate goal of a survey is to generate unbiased estimates of behaviors in the target population. We followed a standard survey research approach to weight the data so as to align the sample more closely with the target population. All of the results presented below and in the appendices are based on weighted data.

Baseline

Baseline targeted population survey data were provided by NORC to the SEIGMA research team with statistical weights that accounted for the survey design, screening rates, completion rates, and post-stratification to the 2015 Massachusetts (MA) population based on four variables (age, gender, race/ethnicity and education). Data from the survey were weighted to account for the stratified survey design (wt1), differential screening rates associated with address characteristics (wt2), response completion rates (wt3), and the number of 18+ household members (wt4). Next, 2015 Census estimates of the MA 18+ population from Public Use Microdata Sample (PUMS) data were used to form 10 raking variables. An iterative raking process was used until marginal weights converged to PUMS totals (sp1wt6). A detailed description of our data weighting procedures for the Springfield baseline target population survey is included Appendix A1.

Follow-up

Follow-up targeted population survey data were provided by NORC to the SEIGMA research team with statistical weights. The weights accounted for the survey design, screening rates, completion rates, and post-stratification to the 2018 Massachusetts (MA) population based on four variables (age, gender, race/ethnicity and education). Data from the survey were weighted to account for the stratified survey design (wt1), differential screening rates associated with address characteristics (wt2), response completion rates (wt3), and the number of 18+ household members (wt4). Next, 2018 Census estimates of the MA 18+ population from Public Use Microdata Sample (PUMS) data were used to form 10 raking variables.[3] An iterative raking process was used until marginal weights converged to PUMS totals (st_wt6). A detailed description of our data weighting procedures for the Springfield Follow-up target population survey is included Appendix A2.

Table 1 on the following page compares key demographic characteristics of the weighted baseline and follow-up targeted population samples along with information about the adult population of the host and surrounding communities (H&SCs) in 2015 and 2019. This comparison is helpful to understand the impact of weighting on the results of the survey.

A comparison of percentages in the weighted column and the Massachusetts 2013 column in Table 1 shows a close match for gender and ethnicity. This is to be expected since these variables were used in the weighting. The comparison of percentages between columns for age is not as close, since the number of age groups used in weighting the sample was smaller than the number of age groups displayed in Table 1.[4] A comparison for education shows that respondents in the FTPS-S had somewhat higher educational attainment compared to the Massachusetts population in 2019. A comparison for income shows that respondents in both the BTPS-S and the FTPS-S had lower annual household incomes compared to the Massachusetts population in 2015 and 2019. These observations suggest that the weighted survey results over-represent adults in higher-education but lower income households in the population.

[3] PUMS 2018 estimates were used because the 2019 estimates were not yet available when the weighting was done.

[4] Four age categories were used in the weighting procedure (18-34, 35-49, 50-64, 65+).

Item non-response was not a major issue in any of the data collection modes. Respondents were allowed to refuse to answer any question or to give a "don't know" response. The percentage of complete responses was extremely high for nearly all items. For interested readers, the response rate for individual questions by data collection mode is shown in Appendix A3. Household income was the only measure that had a non-response rate greater than 10%. Although household income is a candidate for imputation, no imputation was done for this report.

Table 1. Demographics of Springfield Baseline and Follow-up Target Survey Sample

		BTPS (2015)				FTPS (2019)			
		PUMS[5]		SEIGMA		PUMS[6]		SEIGMA	
		%	SE	%	SE	%	SE	%	SE
Gender	Male	46.9	0.9	46.0	2.2	46.9	0.9	46.9	2.2
	Female	53.1	0.9	54.0	2.2	53.1	0.9	53.1	2.2
Age	18-20	6.5	0.4	6.3	1.5	6.4	0.5	9.4	1.8
	21-24	7.3	0.5	5.3	1.1	7.0	0.5	5.5	1.0
	25-34	15.8	0.7	18.1	1.9	16.3	0.7	14.8	1.4
	35-54	32.8	0.9	31.4	2.1	31.0	0.8	31.7	2.1
	55-64	17.5	0.7	18.5	1.5	17.9	0.7	17.1	1.6
	65-79	14.3	0.6	14.6	1.2	16.3	0.6	17.0	1.5
	80+	5.7	0.4	5.8	0.7	5.2	0.4	4.5	0.8
Ethnicity	Hispanic	16.9	0.7	15.8	1.9	18.8	0.8	19.1	1.6
	White alone	72.8	0.9	70.1	2.2	70.7	0.9	67.0	2.0
	Black alone	6.2	0.5	6.3	1.2	6.3	0.5	6.3	1.0
	Asian alone	2.4	0.3	***2.4***	***0.7***	2.7	0.4	2.7	0.8
	Some other race alone	0.3	0.1	***1.0***	***0.4***	0.5	0.1	***2.3***	***0.8***
	Two or more races	1.4	0.2	4.5	0.9	0.9	0.2	2.6	0.7
Education	Less than high school	12.1	0.6	12.6	1.8	13.4	0.7	9.6	1.6
	HS or GED	30.1	0.8	28.7	2.2	28.4	0.8	32.0	2.4
	Some college	30.3	0.8	30.4	1.9	30.3	0.8	27.5	1.7
	BA	15.9	0.6	16.7	1.3	16.6	0.6	19.7	1.4
	Graduate or professional degree	10.4	0.5	9.6	0.8	9.8	0.5	8.3	0.7
	PHD	1.1	0.2	2.0	0.4	1.4	0.2	2.9	0.5
Income	Less than $15,000	9.2	0.6	14.0	1.8	8.6	0.5	14.9	1.8
	$15,000-<$30,000	11.7	0.6	16.2	1.7	12.2	0.6	11.5	1.3
	$30,000-<$50,000	16.6	0.7	21.0	2.0	15.2	0.7	17.8	1.9
	$50,000-<$100,000	29.5	0.8	30.8	2.1	28.2	0.8	31.1	2.1
	$100,000-<$150,000	17.6	0.7	11.9	1.2	20.0	0.7	14.8	1.7
	$150,000 and more	15.3	0.7	6.2	0.9	15.8	0.6	10.0	1.5

Note: Bold, italics indicates estimates are unreliable, relative standard error >30%

Assessing Potential Bias

Research suggests that the main reason for differences in responses in gambling surveys is mode of survey administration, with self-administered questionnaires shown to reduce the potential for bias (R. J. Williams &

[5] Source: Census Bureau, 2015 American Community Survey PUMS
[6] Source: Census Bureau, 2018 American Community Survey PUMS

Volberg, 2009, 2010). The sequence of offering the baseline and follow-up surveys via Web followed by SAQ followed by telephone was intentionally designed to maximize the opportunity for the survey to be self-administered and to reduce the potential for bias. In response to a suggestion from the MGC Research Review Committee, we undertook an analysis to assess whether higher rates of completion by Web and the use of targeted lists in the follow-up survey in 2019 added measurable bias to the results.

These analyses showed that administration mode was significantly related to age, education and income with SAQ and telephone respondents in 2015 and SAQ respondents in 2019[7] significantly more likely to be aged 55 and older compared with Web respondents and significantly less likely to have attended college or graduate school. SAQ respondents in 2019 were significantly less likely than Web respondents to have annual household incomes over $50,000. Since age, education and income are related to gambling attitudes and behavior, these demographic differences are likely responsible for the only apparent mode effect that was identified, with SAQ and telephone completers significantly more likely to believe that gambling was too widely available in Massachusetts in both surveys.

There were no statistically significant differences between respondents who were recruited using targeted lists compared with respondents recruited using ABS in 2019 on any of the key variables. The same result was obtained when the analysis was limited to Black and Hispanic respondents who were recruited via targeted lists compared with ABS. Our conclusion was that the higher rates of completion by Web and the use of targeted lists in 2019 did not add measurable bias to the results.

Data Cleaning and Statistical Analysis

Throughout data collection, SAS programs were run by NORC to identify any errors that occurred in the online or CATI systems. This allowed inconsistencies to be reconciled and for systems or questionnaire errors to be fixed as they occurred. Once data collection was complete, NORC reviewed verbatim responses for several questions that offered an "Other" response category. The verbatim responses were back-coded into existing response categories where appropriate.

After the dataset was received by the SEIGMA research team, skip patterns and outliers were reviewed and a cleaned dataset was created. Using the cleaned data, several additional summative and/or composite variables were created and added to the final dataset.

Statistical analysis of survey data where respondents have unequal weights is more complex than standard statistical analysis due to the need to properly account for the weights in estimating parameters and their variance. Special software and statistics have been developed for such situations. The BTPS-S and FTPS-S data were analyzed using SAS-callable SUDAAN, release 11.0.1. SUDAAN enables appropriate calculation of variance estimations for data from surveys using complex sampling strategies. When exact expressions for the variance were not possible, the Taylor series linearization method was used combined with variance estimation formulas specific to the sample design. Chi-square analysis and other nonparametric techniques were used to test for statistical significance in the sections of this report addressing gambling behavior, problem gambling prevalence and correlates of problem gambling.

Key Informant Interviews

The SEIGMA team conducted interviews from September 2019 to May 2020 with key informants in Springfield. The goal was to gain an on-the-ground understanding from local experts about conditions in Springfield prior to hosting a casino, during the construction of the casino, and while hosting the casino. The SEIGMA team

[7] The number of telephone completers in 2019 (n=15) was too small to affect these analyses.

developed a list of contacts in organizations from Springfield that had hosted events attended by SEIGMA researchers or themselves attended a SEIGMA event, such as a Public Research Day. The focus was on individuals who, through their professional expertise and experience working in the city, could further inform understandings of social conditions within the host community. We requested a single interview from potential key informants by contacting their professional offices by email and/or telephone. If a key informant agreed to an interview, the 60-minute to 90-minute interview was conducted in person, or by Zoom after mid-March 2020. Prior to the interview commencing, formal consent was obtained. Interviews were recorded and were not confidential as officials/representatives spoke in their professional capacity and in their area of expertise.

Interviews were conducted with the following individuals:

- **Rebecca Bishop,** Director, Gambling Prevention Technical Assistance Center, Education Development Center (EDC), May 8, 2020
- **Dr. Stephen Boos,** Medical Director, Family Advocacy Center, Baystate Health Systems, April 7, 2020
- **Jessica Collins,** Executive Director, Public Health Institute of Western Massachusetts, September 23, 2019
- **Amy Gabrila,** Senior GameSense Advisor at MGM Springfield, Massachusetts Council on Gaming and Health, May 14, 2020
- **Chrismery Gonzalez,** Program Lead, Office of Problem Gambling Prevention, Department of Health and Human Services, City of Springfield, September 24, 2019
- **Joesiah Gonzalez,** Director of Youth Services, New North Citizens Council, March 12, 2020
- **Ronn Johnson,** President and CEO, Martin Luther King, Jr. Family Services, April 2, 2020
- **Frank Robinson,** Vice President, Public Health and Community Relations, Baystate Health Systems, January 30, 2020
- **Dr. Jessica Wozniak,** Manager, Clinical Research & Development, Family Advocacy Center, Baystate Health Systems, April 7, 2020

Interviews with key informants covered two major topic areas. The first topic area addressed *Social and Cultural Impacts* and asked how the presence of the casino in Springfield had affected public attitudes, leisure pursuits, interactions with and trust of others (social capital), education, and employment. The second topic area addressed *Public Health Impacts* and asked how the presence of the casino in Springfield had affected families, health and well-being, gambling behavior, rates of problem gambling, bankruptcies, and suicides. At the end of each interview, key informants were asked for recommendations of individuals representing other organizations in Springfield who they felt could provide additional perspective on the topics of the interview. All of the recommended individuals were added to the list of potential key informants and interviews were completed with several of them.

Recordings of the key informant interviews were reviewed independently by two SEIGMA team members and quotes relevant to the topics assessed in the interviews were extracted. In extracting quotes for this report, the team members focused on comments related to gambling attitudes, gambling behavior, problem gambling, and problem gambling services. Quotes extracted by the team members were reviewed by the SEIGMA lead investigator and senior author of this report (Volberg) using a 'matrix' approach to qualitative data analysis developed by the National Centre for Social Research in Britain (Reith & Dobbie, 2013). This approach involves arranging quotes that relate to specific topics in columns with study participants arranged in rows to allow for comparison of quotes across interviews as well as comparison of quotes within interviews. This approach lets researchers conduct both thematic and case-based analysis of qualitative data.

Once quotes had been selected, individuals who had made the statements were contacted and asked to confirm their agreement to the inclusion of specific quotes in the report. Agreement to the inclusion of all the quotes in the report has been confirmed by the relevant key informant.

Reporting

In reporting results, we have used several conventions to make the interpretation of results easier. For example, we adopted the approach used by the National Center for Health Statistics to identify and flag all estimates with a relative standard error (RSE) greater than 30% as not meeting standards for reliability. Standard error (SE) measures the extent to which a survey estimate is likely to deviate from the true value in the population; relative standard error is expressed as a percentage of the survey estimate. Within the report, estimates with RSE greater than 30% are suppressed. In appendices to the report, estimates with RSE greater than 30% are highlighted in red to allow readers to judge these data for themselves. Another measure taken to enhance confidence in the results of the survey was to suppress values in any cells that contained less than five respondents. This was done in both the body of the report and in the appendices.

We have also chosen to present many of our results in graphic form. We have not included the categories of "Don't Know," "Refused," and "Other" in order to make the graphics easier to read. We have included all of the data in tables in the appendices for readers who prefer a tabular format. Throughout the report, we have focused on five major demographic groups (i.e., gender, age, race/ethnicity, education, annual household income). Finally, we discuss differences between groups only when the overall test for group differences is statistically significant based on a chi-square or t-test with alpha of 0.01.[8] The p-values for such tests are presented in the tables in the appendices.

[8] Scholarly convention has long been to use a 0.05 significance level to identify statistically significant findings. However, given the large number of independent statistical tests in the present analysis, the research team deemed that a more conservative 0.01 level was warranted.

Results

In this section of the report, we present descriptive results from the baseline and follow-up targeted surveys in Springfield and surrounding communities. We first examine changes in attitudes toward gambling among residents of these communities, followed by changes in gambling participation and expenditures. We then examine changes in problem gambling prevalence and awareness of problem gambling programs. It is important to emphasize that the targeted surveys in Springfield and surrounding communities were cross-sectional 'snapshots' of gambling and problem gambling at single points in time. This limits our ability to draw any cause and effect conclusions from associations reported between gambling participation, gambling problems, and other variables in these surveys.

Note Regarding Confidence Intervals

In epidemiological research, prevalence is a measure of the number of individuals in the population reporting a behavior or classified with a disorder at one point in time. Prevalence rates are based on samples rather than the entire population. Even when a sample is representative of the population from which it is drawn, an identified value—such as the prevalence rate—is still an estimate and can be different, even if only slightly, from the 'true' value. One important source of uncertainty in generalizing from a sample to the population—sampling error—is generally presented as a measure of the uncertainty around the identified value. This measure is called the confidence interval and it is a gauge of how certain we are that the result we have identified is accurate. The conventional size of the confidence interval is 95% which means that, if a researcher drew 100 samples from the same population, the identified value would fall between the lowest and highest values of the confidence interval 95 times.

Generally speaking, narrower confidence intervals are considered more reliable because the identified value will not be very different in other samples drawn from the same population. As sample size increases, confidence intervals typically narrow. Conversely, as sample size decreases, confidence intervals widen. While the overall size of the samples in both of the Springfield targeted surveys is large, there are some groups in the sample that are quite small. We urge readers to treat estimates based on small groups in the sample with caution and to pay particular attention to the confidence intervals surrounding these estimates.

As a reminder, all of the results presented here and in the appendices are based on weighted data.

Attitudes about Gambling

Before examining gambling participation among residents of Springfield and surrounding communities, it is helpful to consider changes in attitudes toward gambling among residents of these communities. Respondents in both surveys were asked several questions about their views of gambling. Questions assessed respondents' beliefs about legalized gambling in general, the availability of gambling in Massachusetts, the likely impact of expanded gambling in Massachusetts and the likely impact of expanded gambling in their own community, the overall benefit or harm of gambling in society, and the morality of gambling.

Attitudes about Gambling Legalization

Figure 1 shows changes in attitudes about gambling among residents of Springfield and surrounding communities between 2015 and 2019. In both surveys, the majority of residents of Springfield and surrounding communities believed that some forms of gambling should be legal and some should be illegal, with only a minority reporting that all forms should be legal or all forms should be illegal.

Beliefs about Gambling Availability

In both 2015 and 2019, the majority of residents of Springfield and surrounding communities believed that the current availability of gambling in Massachusetts was fine. However, a significantly higher proportion of residents in 2019 (27.7%) believed that the current availability of gambling in Massachusetts was too high compared to 2015 (20.3%).

Perceived Impact of Expanded Gambling in Massachusetts

Between 2015 and 2019, there was a reduction in the proportion of respondents who viewed the impact of expanded gambling in Massachusetts as harmful and an increase in the proportion who viewed the impact as neither beneficial nor harmful. In 2015, 38.8% of residents of Springfield and surrounding communities believed that a casino in Massachusetts would be somewhat or very harmful; in 2019, 21.9% of residents believed this would be harmful. In 2015, 16.8% of residents believed that a casino in Massachusetts would be neither beneficial nor harmful; in 2019, 35.0% believed this would be neither beneficial nor harmful. These changes were statistically significant at the 0.01 level.

Perceived Impact of a New Casino or Slot Parlor in One's Own Community

In 2019 as compared to 2015, a smaller proportion of respondents viewed the impact of having a new casino or slot parlor in their own community as beneficial and a larger proportion of respondents viewed the impact as neither beneficial nor harmful. In 2015, 40.7% of residents of Springfield and surrounding communities believed that a casino in their community would be somewhat or very beneficial and 22.8% believed it would be neither beneficial nor harmful. In 2019, the proportion of residents who believed that a casino in their community would be somewhat or very beneficial had reduced to 32.1% while another 32.0% believed it would be neither beneficial nor harmful.

Attitudes about the Importance of Gambling as a Recreational Activity

When respondents were asked to rate the importance of gambling to them as a recreational activity, only 10.1% of respondents in 2015 and 9.4% of respondents in 2019 indicated that gambling was somewhat or very important to them as a recreational activity. About one quarter of respondents (27.5% in 2015 and 22.7% in 2019) indicated that gambling was not very important to them while the majority of respondents (62.4% in 2015 and 67.9% in 2019) indicated that gambling was not at all important to them as a recreational activity. Overall, these changes were statistically significant at the 0.01 level.

Perceived Benefit or Harm of Gambling to Society

All respondents were asked the following question: "Which best describes your belief about the benefit or harm that gambling has for society?" with possible responses that the benefits somewhat or far outweigh the harm, the benefits are about equal to the harm, and the harm somewhat or far outweighs the benefit. There was a range of opinion among residents of Springfield and surrounding communities concerning the relative harm versus benefit of gambling to society. That said, the majority of residents in 2015 believed the harm outweighed the benefits (51.5%) and a significantly greater proportion of residents in 2019 believed the harm outweighed the benefits (64.1%).

Perceived Morality of Gambling

Lastly, all respondents were asked if they believed gambling is morally wrong. As shown below, there was great uniformity concerning the morality of gambling compared with opinions about the benefit or harm of gambling to society and no significant change in this view between 2015 and 2019. In 2015, 81.8% of the residents of

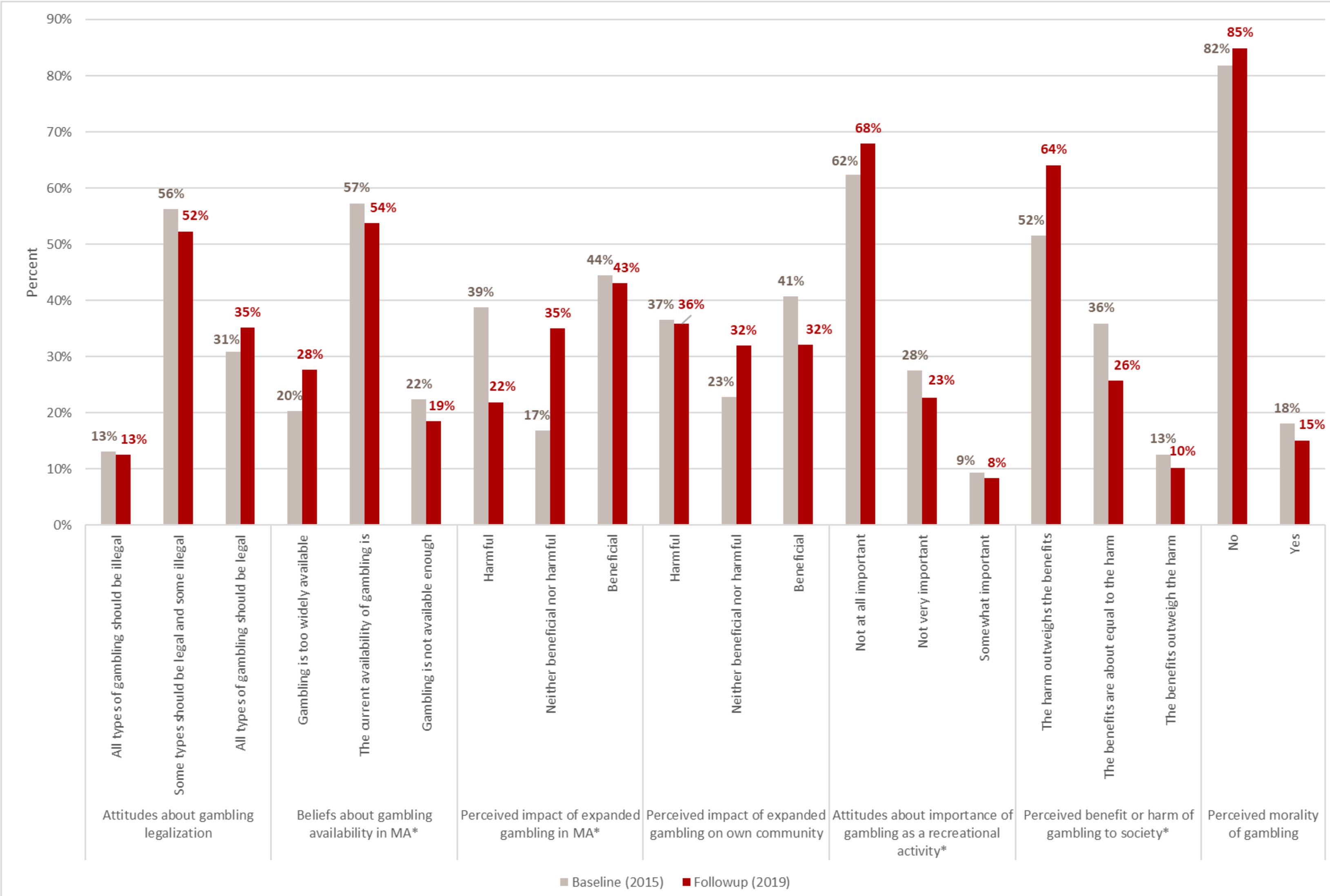
Springfield and surrounding communities did not believe gambling to be morally wrong while in 2019, 84.9% of residents held this view.

To summarize, there were four statistically significant changes in attitudes toward gambling among residents of Springfield and surrounding communities between 2015 and 2019. First, compared to 2015, more residents surveyed in 2019 believed that the availability of gambling in Massachusetts was too high. Second, the majority of residents in both 2015 and 2019 believed that the harm of gambling to society outweighed the benefits with a significantly higher proportion feeling this way in 2019. Third, compared to 2015, more residents viewed the importance of gambling as a recreational activity as "not at all important." Finally, compared to 2015, fewer residents surveyed in 2019 viewed the impact of expanded gambling in Massachusetts as harmful and more residents held a neutral view. Taken together, these changes suggest that overall perceptions of gambling among residents of Springfield and surrounding communities became somewhat more negative but also less polarized between 2015 and 2019.

Several key informants from Springfield commented on attitudes towards gambling prior to and after the opening of the casino. However, the issue of the casino was framed largely in local political terms:

> *"[I] haven't seen a change in attitude towards gambling. Before the casino came to town, saw 'two sides' to the issue... One was the mayor's administration [which] wanted to revitalize the city and bring it back economically… Flip side of that, evangelicals in the city… was on the news a lot protesting that [they] did not want a casino in the city. After it passed, it all died down."* – Joesiah Gonzalez, Director of Youth Services, New North Citizens Council

Figure 1. Change in Attitudes about Gambling from Baseline (2015) to Follow-up (2019)



Estimates that are unreliable (relative standard error >30%) or cell size is 5 or less are excluded from this chart
Asterisk indicates significant change from baseline to follow-up
See Table 2 in Appendix B

Gambling Participation

This section examines changes in gambling participation among residents of Springfield and surrounding communities. To assess the full range of gambling available to Massachusetts residents, the survey included questions about 11 different activities. At the beginning of the survey, all respondents were given the same definition of gambling to assure comprehension and comparability of the results. Respondents were told:

> *We define gambling as betting money or material goods on an event with an uncertain outcome in the hopes of winning additional money or material goods. It includes things such as lottery tickets, scratch tickets, bingo, betting against a friend on a game of skill or chance, betting on horse racing or sports, investing in high risk stocks, etc.*

Gambling participation was assessed by asking about past year frequency of participation in each different type of gambling. In addition to past year frequency of participation, respondents were asked about amounts spent in a typical month on each type of gambling they had done in the past year.

Figure 2 shows changes in overall past year gambling participation as well as past year participation in specific gambling formats among residents of Springfield and surrounding communities between 2015 and 2019. Overall gambling participation in this period changed very little although there was a statistically significant increase in past year participation in daily lottery games from 17.3% in 2015 to 24.1% in 2019. There was no statistically significant change in past year gambling at out-of-state casinos but there was a statistically significant change in any casino gambling in the past year, from 22.5% in 2015 to 36.5% in 2019. "Any casino gambling" includes both out-of-state and Massachusetts casinos; since there were no Massachusetts casinos in 2015, the change in "any casino gambling" was clearly driven by the substantial increase in the proportion of residents of Springfield and surrounding communities in 2019 who gambled at both Massachusetts and out-of-state casinos.

Although the significant increase in participation in daily lottery games may have been due to higher sales, it is more likely that a change in the question wording in 2019 was responsible for the higher reported participation rate. A similar change was observed in Waves 3 and 4 of the MAGIC study where the names of several new daily lottery games (including MassCash, All or Nothing, and the Numbers Game) were added to the list of examples in the questionnaire (R. J. Williams, Zorn, Stanek, Evans, & Volberg, 2020). Obtained rates of participation tend to increase when questions about involvement are posed with more specific detail.

Beyond past year daily lottery participation and overall casino gambling, there were no statistically significant changes in gambling behavior among residents of Springfield and surrounding communities between 2015 and 2019. Past year participation in traditional lottery games, instant lottery games, raffles, sports betting, private betting, horse racing and bingo all remained steady. Additionally, there were no statistically significant changes in maximum frequency of gambling or in the average number of gambling formats engaged with (see Table 3 in Appendix B).

Several key informants commented that there was a well-established gambling culture prior to the casino that was centered around informal community groups and local stores:

> *"The places where local people purchase lottery products (i.e., keno, scratch tickets) are like 'little gambling parlors' and those people frequent those lottery outlets just as much after as they did before the casino was open... this constituency of folks continue to visit the bodegas where they can see their friends, visit with the same set of people, and it becomes a social gathering place. The casino has not*

> *filled that role."* – Frank Robinson, Vice President, Public Health and Community Relations, Baystate Health Systems

Several other key informants commented on the greater accessibility of casino gambling once MGM Springfield opened:

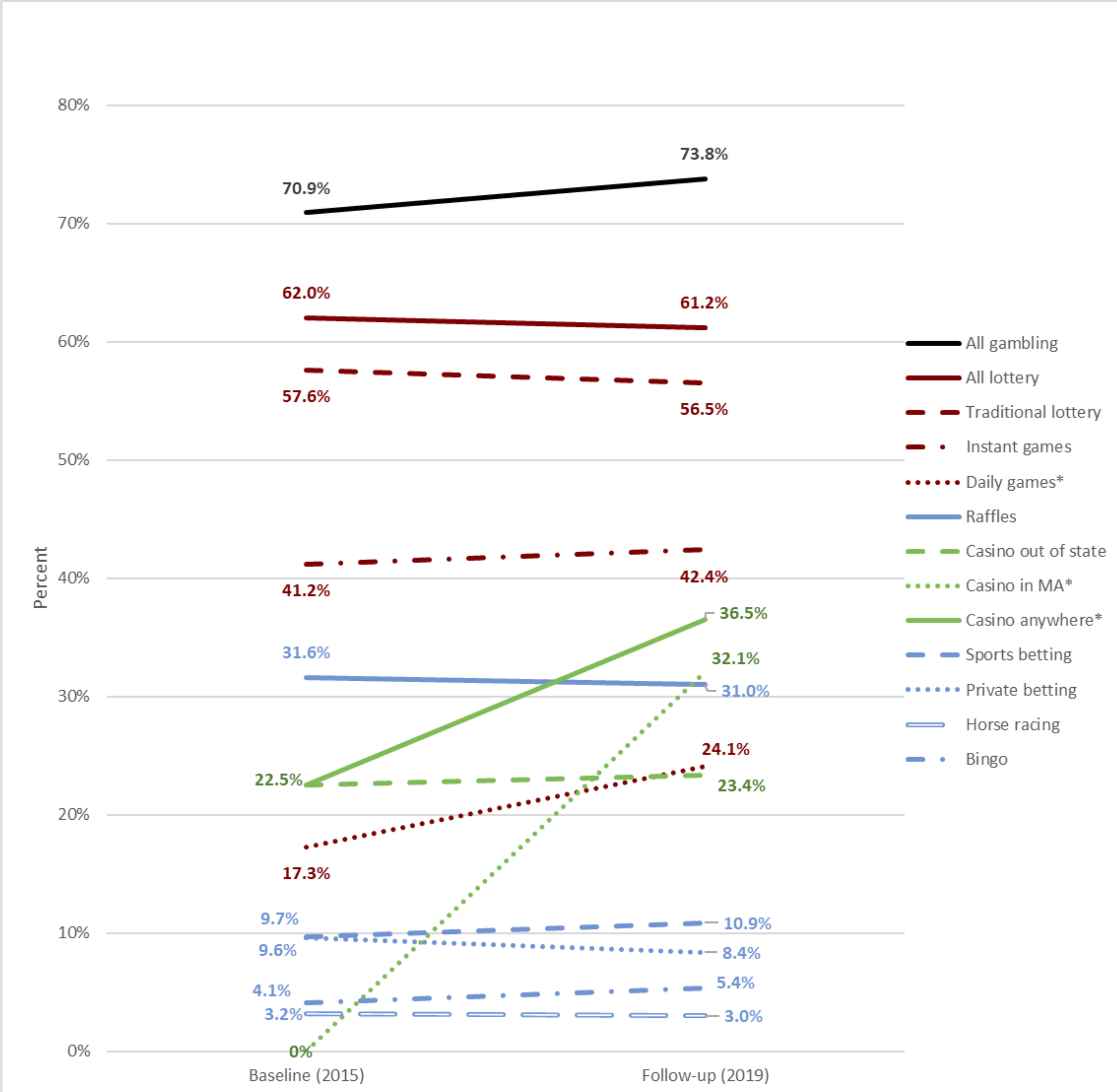
> *"But as I interact with the public, people have expressed they have family members and friends who love gambling, [and that] has increased since the casino has been here because [it is] so accessible."* – Ronn Johnson, President and CEO, Martin Luther King, Jr. Family Services

> *"[The MGM casino] made [gambling] more accessible for folks. Before folks had to get on bus to go to Foxwoods or Mohegan, now [they] can just go after or before work, play a couple machines if they want."* – Joesiah Gonzalez, Director of Youth Services, New North Citizens Council

> *"A lot of them had never gambled in a casino environment before, or had only visited the Connecticut casinos once or twice a year. They have become more regular, not necessarily problematic, and most of them note the proximity of that, [they have] definitely become more aware of the casino environment and maybe frequent it a little more because of that proximity."* – Amy Gabrila, Senior GameSense Advisor at MGM Springfield, Massachusetts Council on Compulsive Gambling

In our report on the Baseline General Population Survey (Volberg et al., 2017: 42-43), we commented on the finding that past year casino participation in Massachusetts as a whole was similar to participation rates in other jurisdictions with well-established casino markets. We hypothesized that if this was the case, the negative impacts of introducing casinos in the Bay State might be less than anticipated since exposure to casino gambling was already high. We address this issue in detail in the *Discussion* section.

Figure 2. Changes in Gambling Participation from Baseline (2015) to Follow-up (2019)



Asterisk indicates significant change from baseline to follow-up
See Table 3 in Appendix B

Gambling Participation by Demographic Group

Given the lack of changes in past year participation in most specific forms of gambling among residents of Springfield and surrounding communities, we felt it was important to explore whether there were changes in overall gambling participation, overall lottery participation, and casino gambling in Massachusetts and out-of-state by specific demographic groups. We looked at differences by gender, age, race/ethnicity, educational attainment, and income level. There were no significant changes in overall gambling or overall lottery participation by gender, race/ethnicity and among adults aged 50 and over. The rate of overall casino gambling was significantly higher in 2019 compared to 2015 among males and females, among Blacks/Hispanics/Asians and Whites/Other,[9] and among individuals aged 50 to 64 and those aged 65 and over. In contrast to most other demographic groups, the rate of overall casino gambling did not change significantly among adults aged 18 to 34 and those aged 35 to 49 between 2015 and 2019. Among adults aged 18 to 34, the rate of overall lottery participation was significantly lower in 2019 compared to 2015 (45.8% in2019 compared to 62.5% in 2015). Tables 4 to 8 in Appendix B present this information in detail.

One key informant commented on casino gambling by people of color in Springfield:

> *"Even with my professional fraternity [which met at MGM Springfield for about a year]… an influential and affluent group, pretty much no one stays behind to gamble, [we] have to walk through the casino but no one plays a game. Education and socio-economic [status] may be a variable, or maybe peer pressure."* – Ronn Johnson, President and CEO, Martin Luther King, Jr. Family Services

Two key informants commented on the relationship between age and gambling participation in Springfield:

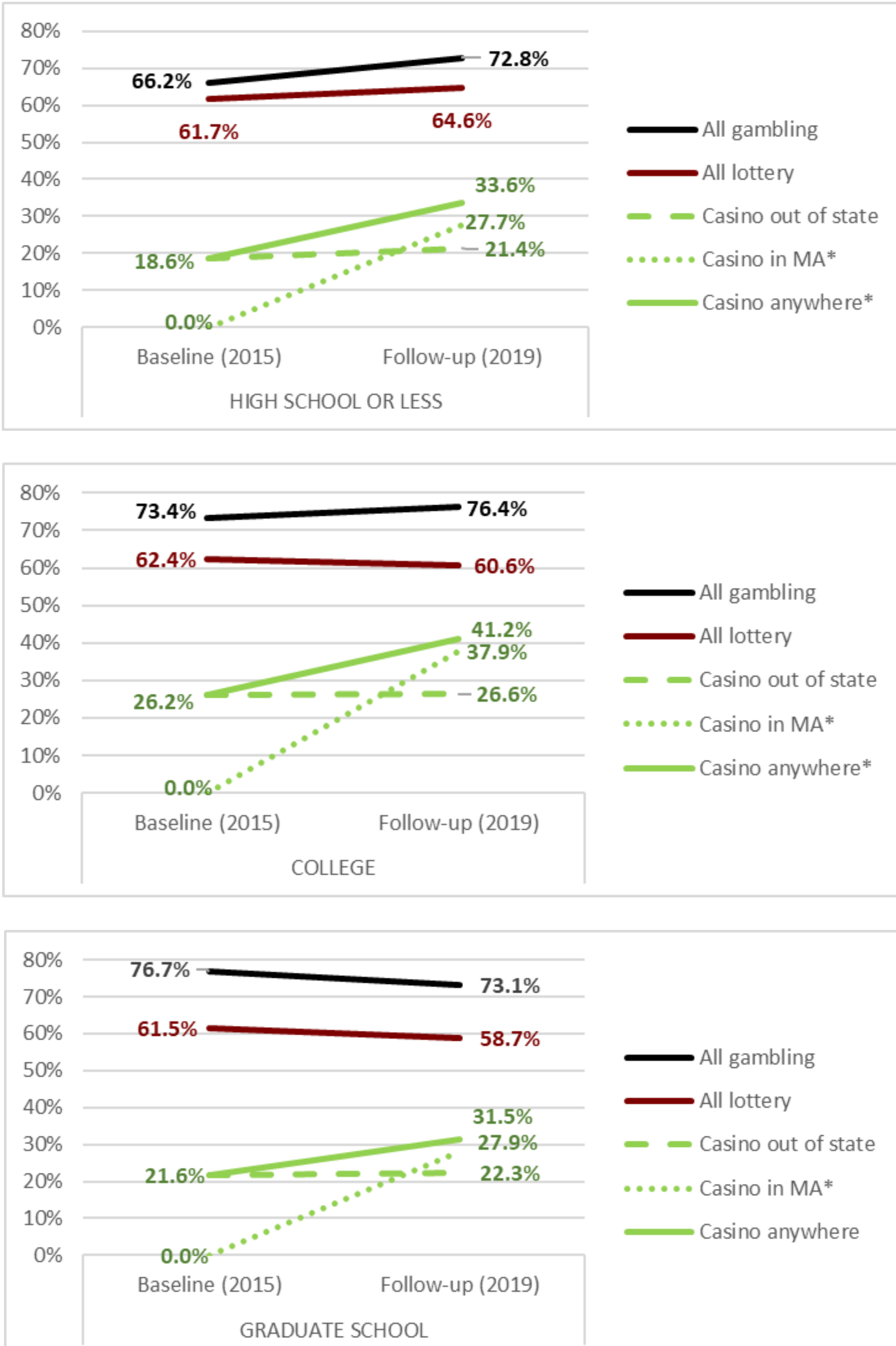
> *"[The] issues we have with our young folks [are] gang enrollments, guns, drugs. Gambling not so much."* – Joesiah Gonzalez, Director of Youth Services, New North Citizens Council

> *"[Gambling] seems to be impacting more the… lower socio-economic status elderly population… We had a recent issue come up in one of the work groups I am a part of, some of the buses from elderly housing in the area [are] delivering individuals to the casino… and having then significant financial difficulties after spending money at the casino as well."* – Jessica Wozniak, Manager, Clinical Research & Development, Family Advocacy Center, Baystate Health Systems

Figure 3 shows that changes in overall casino gambling among residents of Springfield and surrounding communities differed by educational attainment. Rates of overall gambling and overall lottery participation changed very little among residents of Springfield and surrounding communities with different levels of education. The rate of overall casino gambling was significantly higher among those with a high school education or less in 2019 compared to 2015. The rate of overall casino gambling was also significantly higher among those attended or graduated from college. However, the change in the rate of overall casino gambling did not change significantly among those with graduate level education.

[9] The same race/ethnicity groups used in other SEIGMA reports are used in this report to maintain consistency. Given the small proportion of individuals who identified as "Other" (0.09% in the BTPS and 2.16% in the FTPS), grouping these individuals with individuals who identified as White primarily highlights differences between Blacks, Hispanics and Asians, on the one hand, and Whites, on the other.

Figure 3. Changes in Gambling Participation by Education



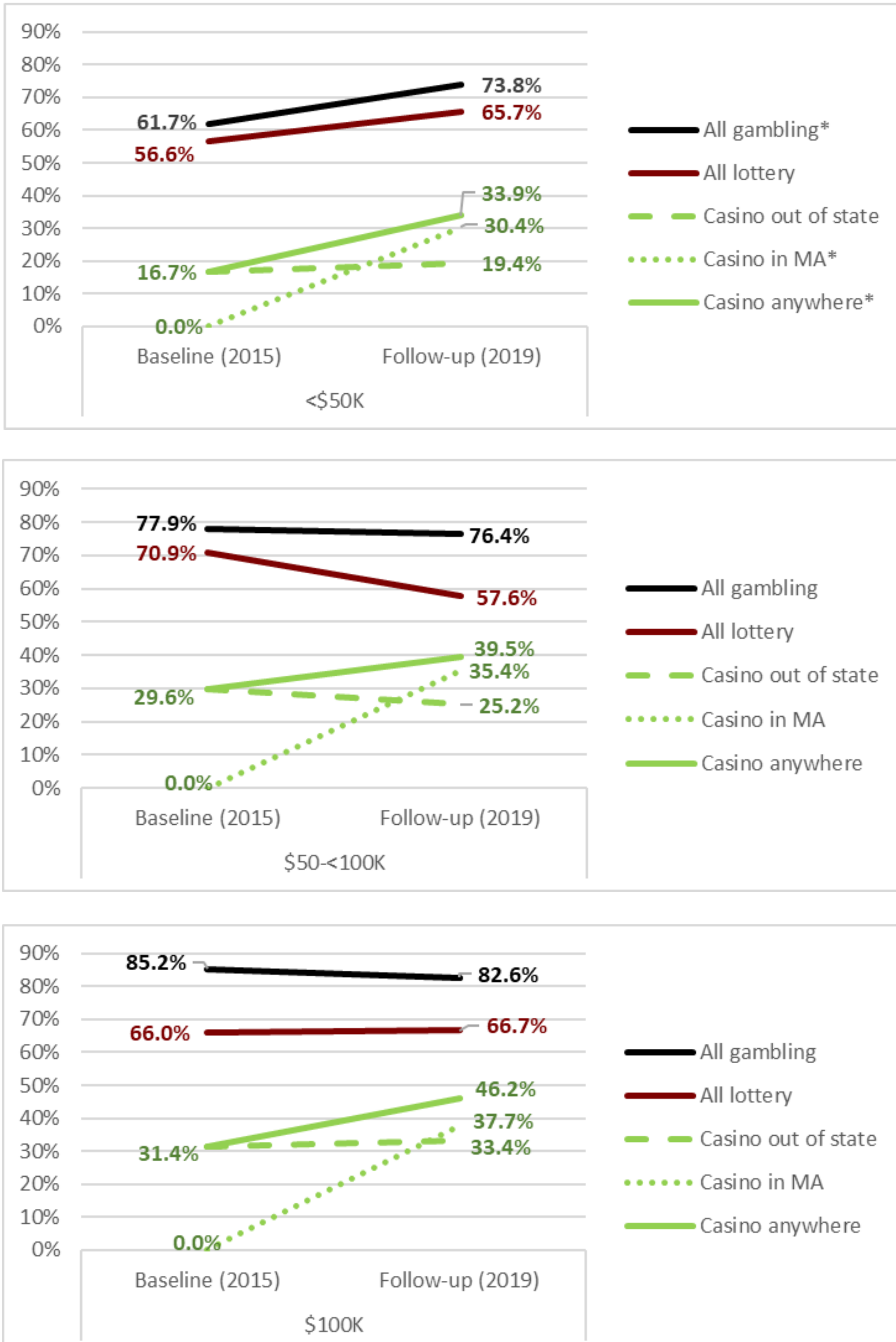
Asterisk indicates significant change from baseline to follow-up
See Table 7 in Appendix B

Figure 4 shows that changes in overall gambling and overall casino gambling among residents of Springfield and surrounding communities differed by income. There were no statistically significant changes in rates of overall gambling, overall lottery participation, or overall casino gambling among residents of Springfield and surrounding communities with annual household incomes of $50,000 or higher. Among residents with annual household incomes lower than $50,000, rates of overall gambling and overall casino gambling were significantly higher in 2019 compared to 2015.

One key informant commented on the relationship between income and casino gambling in Springfield:

> *"I do talk to a lot of lower to middle income, often people of color, that aren't necessarily having a gambling problem, but are gambling more than they were before… a lot of their reasoning is they are trying to better their situation in life. A story I heard, 'I have tried everything in this life to move up, and gambling is my new American Dream. You talk to me about the odds on the slots… but it is just as good as me trying to pull myself up where I am in my situation and do anything legit. The odds aren't that different.' I have heard that story more than once."* – Amy Gabrila, Senior GameSense Advisor at MGM Springfield, Massachusetts Council on Compulsive Gambling

Figure 4. Changes in Gambling Participation by Income



Asterisk indicates significant change from baseline to follow-up
See Table 8 in Appendix B

Gambling Expenditures

Gambling expenditure is an important measure of gambling participation. When survey respondents report accurately, expenditure data are useful to illustrate how much money individuals are spending on different gambling activities. These data, in turn, can be compared to actual and projected revenues, both to validate actual expenditures and to clarify whether revenue projections are accurate. However, surveys have consistently obtained significant underestimates of actual gambling expenditure (Volberg, Gerstein, Christiansen, & Baldridge, 2001; R. J. Williams & Wood, 2007; Wood & Williams, 2007). There are several possible reasons for this lack of correspondence between reported expenditure and actual revenue, including characteristics of different gambling activities, the way in which expenditure questions are asked, respondents' needs to appear socially desirable, and faulty perceptions of wins and losses (Blaszczynski, Dumlao, & Lange, 1997; Volberg et al., 2001; R. J. Williams, Belanger, & Arthur, 2011; Wood & Williams, 2007). Despite these limitations, self-reported expenditure data provide a valuable lens into the relative importance of different gambling activities to the population.

To help understand the impact of the opening of MGM Springfield on gambling expenditures, we compared mean and median total gambling expenditures, total casino gambling expenditures, and total non-casino gambling expenditures in 2015 and 2019. There were no significant changes in total gambling expenditures or total non-casino gambling expenditures between the two time points. There was a statistically significant increase in total in-state and out-of-state casino expenditures although the change in median expenditures was small. Data on gambling expenditures is presented in Table 9 in Appendix B.

Problem Gambling Prevalence

One of the main negative social impacts of expanded gambling availability tends to be an increase in problem gambling (R. J. Williams, Rehm, & Stevens, 2011). As noted previously, prevalence is a measure of the number of individuals in the population with a disorder at one point in time. In epidemiology, prevalence differs from incidence, which is a measure of the number of new cases that arise over a specific period of time. Problem gambling prevalence refers to the percentage of individuals who meet the criteria for problem gambling within the past 12 months. In problem gambling prevalence surveys, individuals are classified on the basis of their responses to a valid and reliable problem gambling instrument.

Measuring Problem Gambling in Massachusetts

Many instruments exist for the population assessment of problem gambling. Worldwide, the most commonly used instruments are the South Oaks Gambling Screen (SOGS) (Lesieur & Blume, 1987), the Canadian Problem Gambling Index (CPGI) (Ferris & Wynne, 2001)[10] and various scales based on the DSM-IV diagnostic criteria for pathological gambling (e.g., Fisher, 2000; Gerstein, Volberg, Harwood, & Christiansen, 1999; Kessler et al., 2008; Petry, Stinson, & Grant, 2005). One or more of these instruments have been used in 95% of adult problem gambling prevalence surveys carried out internationally since 1975 (R. J. Williams, Volberg, & Stevens, 2012).

As explained in detail elsewhere (Volberg et al., 2017), the Problem and Pathological Gambling Measure (PPGM) (R. J. Williams & Volberg, 2014) serves as the primary instrument to assess problem gambling in the SEIGMA study. The PPGM is a 14-item assessment instrument with questions organized into three sections: Problems (7 questions), Impaired Control (4 questions), and Other Issues (3 questions). The instrument employs a 12-month timeframe and recognizes a continuum of gambling across four categories (Recreational, At-Risk, Problem, and Pathological). The PPGM has been field tested and refined with both clinical and general population samples.

[10] Few researchers have used the full 33-item CPGI and the acronym is now commonly used to refer to the shorter, nine-item Problem Gambling Severity Index. We have adopted the same convention in this report.

Internationally, there is widespread agreement that for someone to be classified as a problem gambler there needs to be evidence of both (a) significant negative consequences, and (b) impaired control (Neal, Delfabbro, & O'Neil, 2005). This is made explicit in the PPGM which requires endorsement of one or more items from the Problems section and one or more items from the Impaired Control section to classify an individual as a **Problem Gambler**. In contrast, any pattern of item endorsement that results in a score above a certain threshold is sufficient to be designated as a problem gambler in the SOGS, CPGI, and DSM.[11] Endorsement of several PPGM problems and indices of impaired control is required to classify a person as a **Pathological Gambler**. Endorsement of a problem or impaired control, but not both, typically leads to classification as an **At-Risk Gambler**. This reflects the growing recognition that individuals who become problem gamblers can take a number of different pathways into the disorder (Blaszczynski & Nower, 2002; el-Guebaly et al., 2015; R. J. Williams et al., 2015). Gamblers who do not meet the criteria for At-Risk, Problem, or Pathological Gambling are deemed to be **Recreational Gamblers**. Individuals who have not participated in any of the types of gambling included in the questionnaire are classified as **Non-Gamblers**.

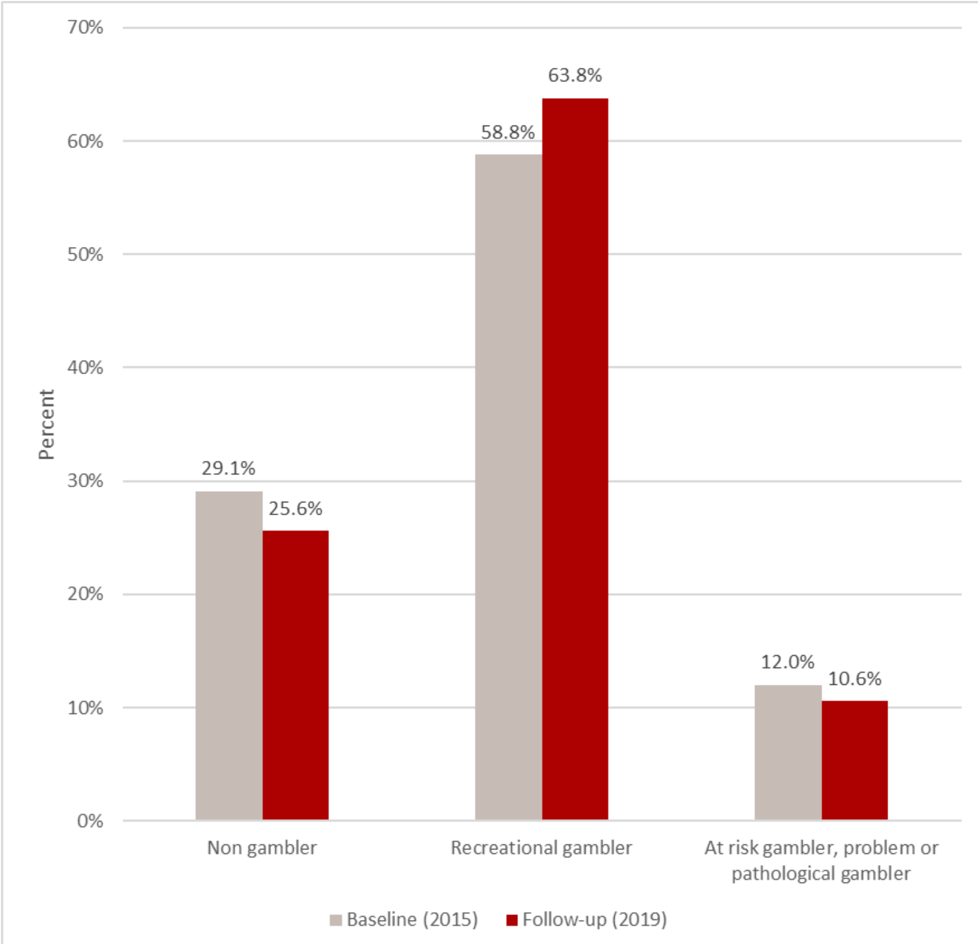
To minimize false positives (i.e., a positive test result that is incorrect), a person has to report gambling at least once a month in the past year to be classified as either a problem or pathological gambler. None of the older problem gambling instruments requires corroborating gambling behavior. To minimize false negatives (i.e., a negative test result that is incorrect) and better identify individuals who have not acknowledged they have a problem, a person can be classified as a problem gambler despite reporting sub-threshold levels of symptomatology if their gambling expenditure and frequency are equal to those of unambiguously identified problem gamblers. While it is well recognized in the addiction field that a significant portion of people with addictions are in denial (Howard et al., 2002; Rinn, Desai, Rosenblatt, & Gastfriend, 2002; Shaffer & Simoneau, 2001), the PPGM is the only gambling instrument designed to identify these individuals.

Figure 5 shows changes in the distribution of gambling types among residents of Springfield and surrounding communities from 2015 to 2019. The figure shows that there was almost no change in the prevalence of at-risk and problem gambling between these two points in time. Although the rate of recreational gambling among residents of Springfield and surrounding communities was higher in 2019 compared to 2015 and the rate of non-gambling was lower, these changes were not statistically significant.

Based on the point estimates and confidence intervals presented in Table 10 in Appendix B, we estimate that between 42,074 (9.3%) and 70,123 (15.5%) residents of Springfield and surrounding communities were at-risk for or experiencing a gambling problem in 2015. We estimate that between 36,421 (8.0%) and 63,281 (13.9%) adult residents of Springfield and surrounding communities were at-risk for or experiencing a gambling problem in 2019. If we consider that each of the individuals at risk for or experiencing a gambling problem is responsible for social and economic impacts that ripple out to their families, friends, employers, and communities, the proportion of the population of Springfield and surrounding communities affected by gambling-related problems is much higher.

[11]All of these problem gambling assessment instruments give each symptom equal weight despite the fact that some items are more serious and/or diagnostically important than others (McCready & Adlaf, 2006; Toce-Gerstein, Gerstein, & Volberg, 2003).

Figure 5. Change in Problem Gambling Prevalence



Estimates that are unreliable (relative standard error >30%) or cell size is 5 or less are excluded from this chart
Asterisk indicates significant change from baseline to follow-up
See Table 10 in Appendix B

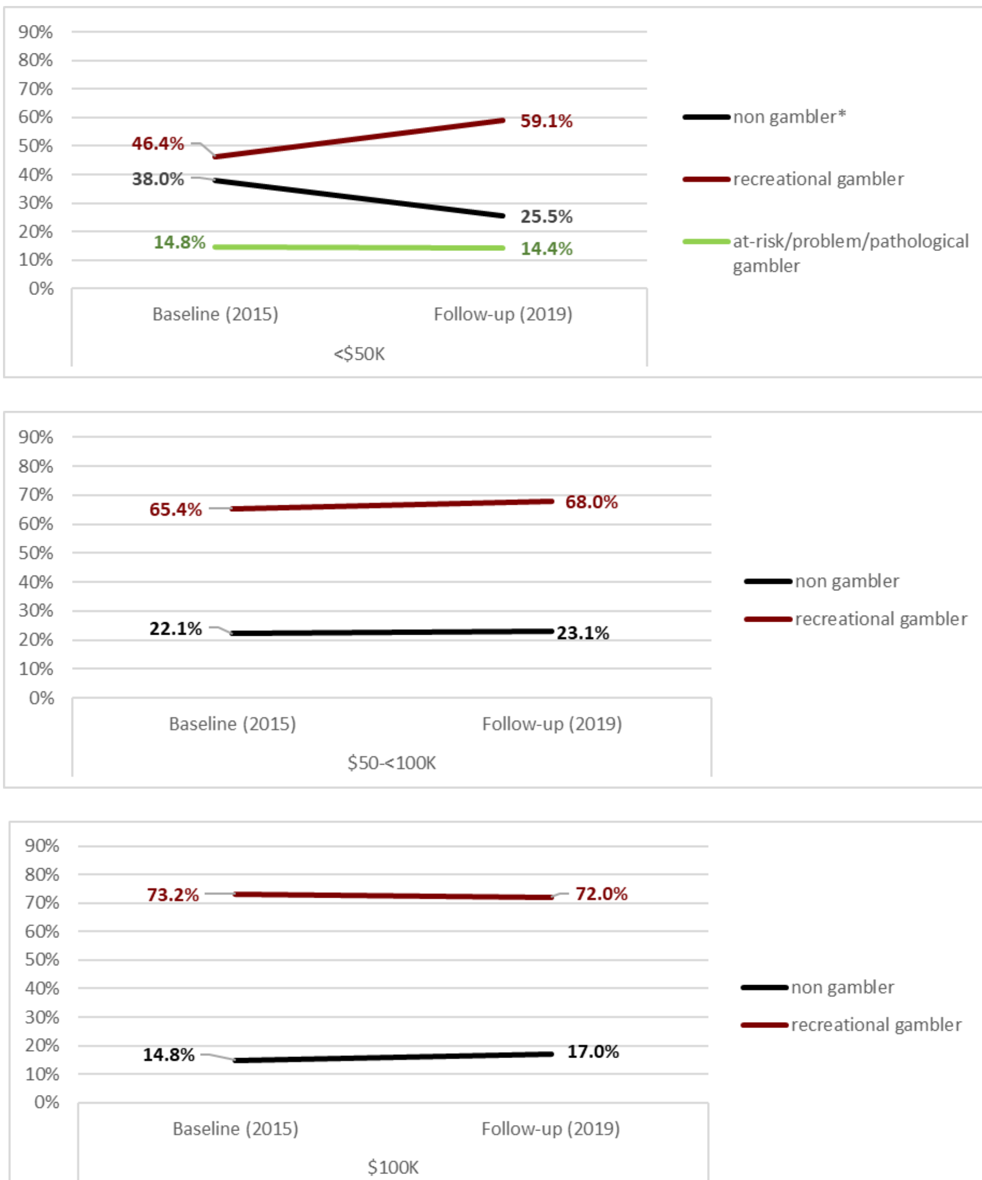
Several key informants commented on the question of whether the rate of problem gambling in Springfield had changed after the opening of the casino:

> *"I belong to tons of groups, am active in church, and [problem gambling] is just not part of the conversation... the only time conversations about this come up are in work, and coordinating efforts with others around accessing public health trust funds, and designing 'upstream, broad interventions,' because the problem gambling just doesn't show up. [I] don't hear people talk about it."* – Frank Robinson, Vice President, Public Health and Community Relations, Baystate Health Systems

Prevalence Rates among Demographic Groups

Problem gambling prevalence rates can be significantly different across important subgroups in the population. Because confidence intervals around prevalence estimates can be large, comparisons between these groups must be interpreted with caution. Nevertheless, this information is important in helping target public health efforts toward groups in the population that are most in need of help. Tables 10-15 in Appendix B present information about the distribution of gambling types among different groups of residents of Springfield and surrounding communities in 2015 and 2019. The only statistically significant change was a lower rate of non-gambling (accompanied by a substantial but not significantly higher rate of recreational gambling) among residents of Springfield and surrounding communities with annual household incomes under $50,000.

Figure 6. Changes in PPGM by Income



Estimates that are unreliable (relative standard error >30%) or cell size is 5 or less are excluded from this chart.
Asterisk indicates significant change from baseline to follow-up
See Table 15 in Appendix B

Awareness of Problem Gambling Programs

Previous research has found that over time, many individuals experiencing gambling problems recover without the aid of professional treatment. Indeed, the literature indicates that the number of people who have recovered on their own may greatly exceed the number of people who ever seek treatment (Castellani, 2000; Hodgins, Currie, el-Guebaly, & Peden, 2004; Korn & Shaffer, 1999). This literature suggests that the behavior of those experiencing gambling problems may be more susceptible to change than was previously thought. This literature also highlights the importance of increasing public awareness and developing brief, targeted interventions to prompt changes in attitudes and behavior among individuals experiencing mild or moderate difficulties to reduce their progression toward more severe gambling-related problems.

In this section, we present information about changes in awareness of and access to problem gambling services among residents of Springfield and surrounding communities. This information is important to understand the extent of general and targeted awareness and prevention programs and to assess the effectiveness of strategies that have been developed to provide help to individuals and groups affected by gambling-related problems.

Awareness of Problem Gambling Prevention Efforts

All of the respondents in both surveys were asked whether they had seen or heard any media campaigns to prevent problem gambling in Massachusetts in the past 12 months. Respondents were also asked whether they were aware of any programs to prevent problem gambling offered in their schools, workplaces, or communities in the past 12 months. Finally, respondents who were aware of a problem gambling program or campaign were asked whether they had participated in any of these programs or campaigns.

As shown in Figure 7, there was a statistically significant reduction in the proportion of residents of Springfield and surrounding communities who indicated that they were aware of media campaigns to prevent problem gambling in Massachusetts in the past year between 2015 and 2019. Almost half of the respondents in the 2015 survey (47.9%) were aware of problem gambling prevention media campaigns in 2015 but the proportion dropped to 32.1% in 2019.

To provide context for this observed change, we spoke with Marlene Warner, Executive Director of the Massachusetts Council on Gaming and Health (formerly the Massachusetts Council on Compulsive Gambling). The Massachusetts Council ran media campaigns addressing problem gambling throughout Massachusetts for many years in coordination with the Massachusetts Department of Public Health. This ended in 2016 when the DPH Office of Problem Gambling Services was established and oversight of problem gambling services in Massachusetts was centralized. It is possible that changes in the administration of problem gambling services in Massachusetts, along with the end of heated public discussion of the casino issue in Western Massachusetts, contributed to these changes. For further context, the lead author of this report (Volberg) travelled through Springfield on I-91 twice a month between January 2013 and July 2017 for personal family reasons. As a gambling researcher, it was notable that billboards advertising the problem gambling helpline in Massachusetts were highly visible on this major travel corridor until late in 2015 when most were replaced with advertising for Foxwoods and Mohegan Sun, the two tribal casinos in Connecticut.

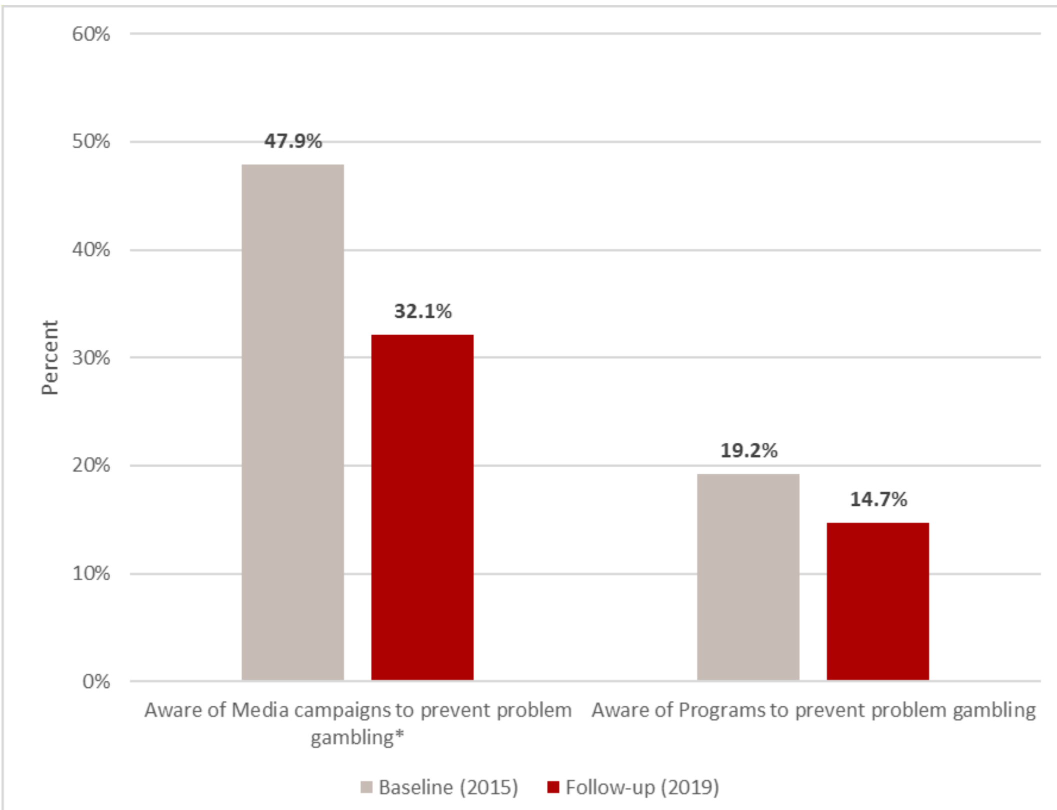
Awareness of problem gambling prevention programs other than media campaigns was lower than awareness of media campaigns. Although there was a reduction in the proportion of residents of Springfield and surrounding communities who indicated that they were aware of programs other than media campaigns from 19.2% in 2015 to 14.7% in 2019, this change was not statistically significant at the 0.01 level.

Two key informants mentioned problem gambling prevention initiatives that have been implemented in the casino or throughout the City of Springfield:

"Based on a casino coming to western Massachusetts, [there have] been more conversations raising awareness about problem gambling—what it looks like, and how [the casino] impacts the public health infrastructure." – Chrismery Gonzalez, Program Lead, Office of Problem Gambling Prevention, Department of Health and Human Services, City of Springfield

"[In relation to Voluntary Self-Exclusion] compared to PPC, I think [there are] more at MGM. Not crazy more, but more. That could be proximity… [and] table games and poker [are a] new segment as well. Maybe Play My Way plays a role down [at PPC], we don't have that here yet." – Amy Gabrila, Senior GameSense Advisor at MGM Springfield, Massachusetts Council on Compulsive Gambling

Figure 7. Change in Awareness of Problem Gambling Programs



Asterisk indicates significant change from baseline to follow-up
See Table 16 in Appendix B

Use of Problem Gambling Treatment Services

Based on publicly available information on the Gamblers Anonymous and Gam-Anon websites, there were three Gamblers Anonymous meetings in the Springfield area and two Gam-Anon meetings as of August 21, 2020 (www.gamblersanonymous.org/ga/locations; www.gam-anon.org/meeting-directory). Two of the GA meetings are held at First Church of Christ in Longmeadow, MA and one takes place at Orchard Covenant Church in Indian Orchard, MA. One of the Gam-Anon meetings is held at Evangelical Covenant Church in Indian Orchard, MA and the other is held at First Church of Christ in Longmeadow, MA.

In response to a query from the SEIGMA team, the office manager at Orchard Covenant Church in Indian Orchard responded by email as follows:

“The GA groups (Gamblers' Anonymous and Gam-Anon) that meet here at our church have been meeting for at least five years, possibly more … The groups have not been meeting in our facility since March”

In response to a follow-up question regarding the impact of the opening of MGM Springfield on the number of group meetings or attendance, she responded:

“I know it is an active group, chomping at the bit to get back to in person meetings. But I … do not know the answer to your question.”

In response to a separate query from the SEIGMA team, the administrator of First Church of Christ responded in a telephone call on August 21, 2020 that the weekly Gamblers Anonymous meetings have been taking place at the location for about 30 years and that, before the pandemic, the meetings usually had 25 to 30 participants per meeting. Since the pandemic, the meetings have been limited to 10 participants although the members are also trying to meet using Zoom and telephone calls. There are additional 12-step meetings in other churches in the area, including two in Springfield and others in East Longmeadow and other surrounding communities. In response to a specific question, the administrator had not noticed any changes in the number of attendees nor was she aware of any new cases, relapses or exacerbations of existing difficulties in the wake of the opening of MGM Springfield.

One key informant commented on the availability of services for people experiencing gambling problems in Springfield:

“My understanding from colleagues that focus on problem gambling treatment, yes, it has significantly impacted. [There has been an] increase of referrals for a lot of the community based mental health clinics around that… and from our understanding there has been a significant increase in those clinics that are specializing in problem gambling.” – Jessica Wozniak, Manager, Clinical Research & Development, Family Advocacy Center, Baystate Health Systems

Discussion

The relationship between proximity, and thereby availability, of gambling venues and the likelihood of experiencing gambling problems has long been debated. The availability, or exposure, theory suggests that an increase in the availability of gambling venues to a population would lead to an increase in gambling-related harms, particularly gambling problems (M. Abbott & Volberg, 1999; Raylu & Oei, 2002; Room, 2005). Alternatively, the adaptation theory suggests that there will be an initial increase in gambling-related problems upon exposure of a resident population to a new gambling venue with the effects abating over time (M. Abbott, 2005, 2006; M. Abbott, Williams, & Volberg, 1999; Shaffer, 2005; Shaffer, Hall, & Vander Bilt, 1997).

Several cross-sectional studies have found that increased exposure to gambling venues is followed by an increase in the prevalence of problem gambling. The *Gambling Impact and Behavior Study* published in 1999 in the U.S. found that the availability of a casino within 50 miles was associated with a higher prevalence rate of problem and pathological gambling (D. Gerstein et al., 1999). Another national U.S. study conducted by Welte and colleagues (2004) found a positive correlation between casino proximity and problem gambling prevalence for those living within 10 miles of a casino. A follow-up study confirmed that individuals who lived closer to casinos had increased gambling participation, and higher rates of problem gambling, than those who did not live in close proximity to a casino (Welte, Barnes, Tidwell, Hoffman, & Wieczorek, 2016). Pearce and colleagues (2008) found that in New Zealand, living near a casino conferred a statistically significant risk for experiencing gambling problems. By looking at casino distances of less than 200 kilometers, Rush and colleagues (2007) found proximity to gambling venues was important in predicting the risk of problem gambling in the exposed population.

However, other cross-sectional studies have found no relationship between casino proximity and prevalence of problem gambling, lending support to the adaptation theory. Sevigny and colleagues (2008) observed a link between gambling venue proximity and gambling participation and expenditure but found no correlation with problem gambling prevalence. Longitudinal studies have found mixed results. Room and colleagues (1999) in Niagara Falls, Canada found an increase in gambling problems in local residents a year after the opening of a new gambling venue. Although Jacques and colleagues (2000) found an increase in the prevalence of at-risk and probable pathological gambling one year after the opening of a local new gambling venue in Quebec Province, Canada, these increases were not maintained at 2- and 4-year follow-ups (Jacques & Ladouceur, 2006). The relationship between casino proximity and gambling problems has never been examined in the northeastern United States, let alone Massachusetts.

Casinos were first introduced into the northeastern United States in the late 1970s, with their legalization in Atlantic City, New Jersey. Expansion continued in the 1990s with the introduction of casinos in Connecticut, Rhode Island, and New York State. Many additional venues have been added since that time. By 2014, the only northeastern states without casinos were Vermont and Massachusetts. This changed in June 2015 with the opening of Plainridge Park Casino, followed by MGM Springfield in August 2018, and Encore Boston Harbor in June 2019.

In a meta-analysis of 34 surveys completed in the Australian states and territories and in New Zealand, Storer and colleagues (2009) noted that decreases in problem gambling prevalence can occur due to a reduction in incidence or problem duration. They argued that a variety of factors, at both the individual and community levels, are likely to influence incidence and problem duration, including natural recovery or professional intervention (at the individual level), and adjusting to the novelty of gambling opportunities or increasing awareness of potential harms (at the community level). They further noted that a decrease in problem gambling prevalence over time could be due to population adaptation in the form of 'natural selection,' with unsuccessful individuals removed from the problem gambling 'pool' for a variety of reasons, including emigration and death.

These different aspects of adaptation suggest quite different policy approaches, with prevention and early intervention more likely to be beneficial in cases where adaptation is taking place at the individual and community level but with stronger measures related to limiting or reducing electronic gambling machine (EGM) density and concentration more likely to be helpful in cases where population adaptation is occurring.

Replication surveys that examine changes in problem gambling prevalence in the same jurisdiction over time provide a direct test of exposure versus adaptation given that gambling availability has generally increased in most jurisdictions over the past 30 years. Replication surveys have been conducted in many jurisdictions and Williams and colleagues (Volberg & Williams, 2014; R. J. Williams, Volberg, & Stevens, 2011) were able to use this body of research to examine changes in problem gambling prevalence over time. This analysis clearly showed that problem gambling prevalence rates in most jurisdictions have tended to decline relative to earlier rates. However, the analysis also showed that the decline has been more dramatic for some jurisdictions (i.e., Canada) relative to others; that the decline started at different times in different jurisdictions (in the late 1990s for Canada and the United States versus the early 2000s for Australia and other nations); and that, in most cases, problem gambling rates *increased* prior to their decline. Problem gambling prevalence rates peaked in the mid to late 1990s for North America and in the early 2000s for Australia and other nations, roughly coincidental with the periods of most rapid introduction and expansion of casino and EGM gambling.

These replication studies support both the contention that increased gambling availability is related to increased problem gambling as well as the contention that populations tend to adapt over time. However, echoing the sentiments of Storer and colleagues (2009), the mechanisms involved in decreasing problem gambling prevalence are probably quite complex. They likely include greater public awareness of the potential harms of gambling, decreased participation once novelty has worn off, increased government and industry efforts to provide gambling more safely, expanding services for those experiencing gambling problems, increased age of the population, and an outflow of problem gambling cases due to severe personal or financial crisis, criminal charges, or, in extreme cases, suicide.

Our findings from the surveys carried out in Springfield and surrounding communities, like our findings from surveys carried out in Plainville and surrounding communities (SEIGMA Research Team, 2018), suggest that the Massachusetts population is far from naïve when it comes to casino gambling opportunities. States surrounding Massachusetts, including Rhode Island, Connecticut, and New York, have had casino gambling for decades prior to the introduction of casino gambling in Massachusetts in 2015. Following this initial exposure, any effects may have abated over time and led to the observed adaptation, even in a population that has experienced recent local gambling expansions. Since problem gambling prevalence in Massachusetts is unknown for the decades prior to 2013, when the SEIGMA Baseline General Population Survey was carried out, one is left to speculate as to an explanation for not observing an increase in gambling problems in the present study. In our view, the findings from this study suggest that population adaptation has already occurred as no increased risk of harms associated with casino gambling have been identified among this previously exposed population.

Aside from this indication that prior exposure to casino gambling for more than two decades appears to have resulted in little measurable impact of the introduction of casino venues within the state, there are additional factors that may have contributed to this perceived adaptation. An increase in public awareness, through media or public health campaigns, may have raised awareness of the potential harms that can occur with increased gambling participation and, subsequently, may have reduced involvement by at-risk individuals. The expansion of treatment services for those individuals who do experience gambling problems may have contributed to increased rates of recovery and fewer relapses. Regulatory or industry measures instituted to curtail gambling harms and increase participant safety, such as casino self-exclusion programs or the Massachusetts Gaming Commission's GameSense program, may also have prevented some at-risk individuals from developing gambling problems.

While the overall results of the surveys in Springfield and surrounding communities are reassuring, there are concerns related to how specific demographic groups in the region may be affected in the future. These groups include individuals with lower education attainment and individuals with annual household incomes under $50,000. As a reminder, the rate of overall casino gambling was significantly higher in 2019 compared to 2015 among those with less than a graduate level education. Among those with annual household incomes under $50,000, the rate of overall gambling as well as the rate of overall casino gambling were significantly higher in 2019 compared to 2015. Also among those with annual household incomes under $50,000, the rate of non-gambling was significantly lower in 2019 compared to 2015. Taken together, these results suggest that there are groups in Springfield and surrounding communities that may be particularly vulnerable to experiencing gambling harms or developing gambling problems because the location of MGM Springfield has made it easier for them to engage in a type of gambling with which they have had relatively little experience in the past. It will be important to direct prevention and treatment resources toward these groups going forward as well as to assess at-risk and problem gambling rates in these groups in the future.

Limitations

There are some limitations to the Baseline and Follow-up Targeted Population Surveys in Springfield and surrounding communities. One important limitation relates to the response rates attained in the surveys, particularly in the Follow-up survey. Survey response rates in developed countries have fallen precipitously in recent years; this increases the likelihood that participants differ from non-participants in some important and systematic way, making the sample non-representative. While this does not always occur (Curtin, 2000; Groves et al., 2006; Keeter, Miller, Kohut, Groves, & Presser, 2000), the risk is always present and tends to increase as a function of the degree of non-response. While we attempted to minimize systematic bias by introducing the study as a survey of "health and recreation," the response rates for both targeted surveys in Springfield and surrounding communities were lower than desirable and, as a consequence, generalization of the results should be undertaken with care.

Another limitation is that the survey was restricted to adults living in households—the sample did not include adults living in group quarters, incarcerated individuals, or homeless individuals. Although rates of problem gambling tend to be very high in these groups, they represent only small proportions of the total population and research has shown that their inclusion is unlikely to affect the overall prevalence rate (M. W. Abbott & Volberg, 2006; R. J. Williams & Volberg, 2010).

A third limitation is that the questionnaire was translated into Spanish but not into other languages. This decision was informed by the fact that the majority of non-English-speaking households in the City of Springfield (and presumably in several of the nearby surrounding communities) are Spanish speaking. Data from the 2018 American Community Survey indicates that Spanish is spoken in 36.8% of households in Springfield while other languages, including Other Indo-European Languages, Asian/Pacific Island Languages, and Other Languages are spoken in 6.9% of households in Springfield. By not providing for surveys in additional languages, we were unable to include such individuals in our sample. However, it is our belief that alternate research strategies are needed to fully explore the role of gambling in a variety of small but important cultural communities, including Asians and South Asians as well as immigrant and refugee communities.

A fourth limitation relates to the small size of several subgroups in the sample such that the prevalence rates of at-risk and problem gambling in these groups are associated with large confidence intervals and should be viewed with caution. It is important to emphasize that while the true value of the estimate of the prevalence of at-risk and problem gambling in these subgroups is somewhere between confidence intervals, the confidence intervals themselves are not unreliable. Any estimates presented in the body of this report can be considered reliable since our convention is to suppress any estimates with relative standard error (RSE) greater than 30%. Estimates in the tables in the appendices with RSE greater than 30% have been flagged and should be viewed with caution since they may be unreliable based on the RSE.

Finally, it is important to emphasize that, like other prevalence surveys, the targeted surveys in Springfield and surrounding communities are cross-sectional ‘snapshots’ of gambling and problem gambling at single points in time. This limits our ability to draw any cause and effect conclusions from associations reported between gambling participation, gambling problems, and other variables in these surveys.

Future Directions

In the coming years, we look forward to continuing the SEIGMA project to assist in understanding the social and economic impacts of gambling in Massachusetts and minimizing and mitigating the negative impacts. With regard to primary data collection, we anticipate conducting several online focus groups with representatives of key stakeholder organizations in Springfield and surrounding communities in the first half of 2021. We also anticipate fielding a large state-wide follow-up general population survey in the Fall of 2021 as well as a first patron survey at Encore Boston Harbor in the Winter of 2022. All of this work, along with efforts by the SEIGMA Economic team, will feed into an integrated report on the social and economic impacts of all three of the casinos in Massachusetts that we anticipate will be published sometime in 2023.

References

Abbott, M. (2005). Disabling the public interest: gambling strategies and policies for Britain: a comment on Orford 2005. *Addiction, 100*(9), 1233-1235; discussion 1235. doi:10.1111/j.1360-0443.2005.01198.x

Abbott, M. (2006). Do EGMs and problem gambling go together like a horse and carriage? *Gambl Res, 18*.

Abbott, M., & Volberg, R. (1999). *Gambling and Problem Gambling in the Community: An International Overview and Critique.* Retrieved from Wellington, NZ:

Abbott, M., Williams, M., & Volberg, R. A. (1999). *Seven years on: A follow-up study of frequent and problem gamblers living in the community.* Retrieved from Wellington: https://www.dia.govt.nz/pubforms.nsf/URL/report2.pdf/$file/report2.pdf

Abbott, M. W., & Volberg, R. A. (2006). The measurement of adult problem and pathological gambling. *International Gambling Studies, 6*(2), 175-200.

Blaszczynski, A., Dumlao, V., & Lange, M. (1997). How much do you spend gambling? Ambiguities in survey questionnaire items. *Journal of Gambling Studies, 13*(3), 237-252.

Blaszczynski, A., & Nower, L. (2002). A pathways model of problem and pathological gambling. *Addiction, 9*, 487-499.

Castellani, B. (2000). *Pathological gambling: The making of a medical problem.* Albany, NY: State University of New York Press.

Curtin, R., Presser, S., & Singer, E. (2000). The effects of response rate changes on the Index of Consumer Sentiment. *Public Opinion Quarterly, 64*, 413-428.

el-Guebaly, N., Casey, D. M., Currie, S., Hodgins, D. C., Schopflocher, D., Smith, G. J., & Williams, R. J. (2015). *The Leisure, Lifestyle, and Lifecycle Project (LLLP): A longitudinal study of gambling in Alberta*. Retrieved from Edmonton: http://dspace.ucalgary.ca/bitstream/1880/50377/1/LLLP_Final_Report_Feb21_2015_V4.pdf

Ferris, J., & Wynne, H. (2001). *The Canadian Problem Gambling Index: Final report*. Retrieved from Ottawa: http://www.ccgr.ca/wp-content/uploads/2013/03/CPGI-Final-Report-English.pdf

Fisher, S. E. (2000). Measuring the prevalence of sector-specific problem gambling: A study of casino patrons. *Journal of Gambling Studies, 16*(1), 25-51.

Gerstein, D., Hoffman, J. H., Larison, C., Engelmann, L., Murphy, S., Palmer, A., . . . Sinclair, S. (1999). *Gambling Impact and Behavior Study: Report to the National Gambling Impact Study Commission.* Retrieved from

Gerstein, D. R., Volberg, R. A., Harwood, H., & Christiansen, E. M. (1999). *Gambling impact and behavior study: Report to the National Gambling Impact Study Commission*. Retrieved from Chicago, IL: http://www3.norc.org/projects/Gambling+Impact+Behavior+Study.htm

Groves, R. M., Couper, M. P., Presser, S., Singer, E., Tourangeau, R., Piani Acosta, G., & Nelson, L. (2006). Experiments in producting nonresponse bias. *Public Opinion Quarterly, 70*(5), 720-736.

Hodgins, D. C., Currie, S., el-Guebaly, N., & Peden, N. (2004). Brief motivational treatment for problem gambling: A 24-month follow-up. *Psychology of Addictive Behaviors, 18*(3), 293-296.

Howard, M., McMillen, C., Nower, L., Elze, D., Edmond, T., & Bricout, J. (2002). Denial in addiction: toward an integrated stage and proces model-qualitative findings. *Journal of Psychoactive Drugs, 34*(4), 371-382.

Jacques, C., & Ladouceur, R. (2006). A prospective study of the impact of opening a casino on gambling behaviours: 2- and 4-year follow-ups. *Can J Psychiatry, 51*(12), 764-773. doi:10.1177/070674370605101206

Jacques, C., Ladouceur, R., & Ferland, F. (2000). Impact of availability on gambling: a longitudinal study. *Can J Psychiatry, 45*(9), 810-815. doi:10.1177/070674370004500904

Keeter, S., Miller, C., Kohut, A., Groves, R. M., & Presser, S. (2000). Consequences of reducing nonresponse in a national telephone survey. *Public Opinion Quarterly, 64*, 125-148.

Kessler, R. C., Hwang, I., LaBrie, R. A., Petukhova, M., Sampson, N. A., Winters, K. C., & Shaffer, H. J. (2008). The prevalence and correlates of DSM-IV Pathological Gambling in the National Comorbidity Survey Replication. *Psychological Medicine, 38*(9), 1351-1360.

Korn, D. A., & Shaffer, H. J. (1999). Gambling and the health of the public: Adopting a public health perspective. *Journal of Gambling Studies, 15*(4), 289-365.

Lesieur, H. R., & Blume, S. B. (1987). The South Oaks Gambling Screen (SOGS): A new instrument for the identification of pathological gamblers. *American Journal of Psychiatry, 144*(9), 1184-1188.

McCready, J., & Adlaf, E. (2006). *Performance and enhancement of the Canadian Problem Gambling Index (CPGI): Report and recommendations*. Retrieved from Guelph: http://www.gamblingresearch.org/

Neal, P., Delfabbro, P. H., & O'Neil, M. (2005). *Problem gambling and harm: Towards a national definition*. Retrieved from Adelaide: www.gamblingresearch.org.au

Pearce, J., Mason, K., Hiscock, R., & Day, P. (2008). A national study of neighbourhood access to gambling opportunities and individual gambling behaviour. *J Epidemiol Community Health, 62*(10), 862-868. doi:10.1136/jech.2007.068114

Petry, N. M., Stinson, F. S., & Grant, B. F. (2005). Comorbidity of DSM-IV pathological gambling and other psychiatric disorders: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Journal of Clinical Psychiatry, 66*, 564-574.

Raylu, N., & Oei, T. P. S. (2002). Pathological gambling: A comprehensive review. *Clinical Psychology Review, 22*(7), 1009-1061. doi:10.1016/S0272-7358(02)00101-0

Reith, G., & Dobbie, F. (2013). Gambling careers: A longitudinal, qualitative study of gambling behavior. *Addiction Research & Theory, 21*(5), 376-390.

Rinn, W., Desai, N., Rosenblatt, H., & Gastfriend, D. R. (2002). Addiction denial and cognitive dysfunction: A preliminary investigation. *Journal of Neuropsychiatry, 14*(1), 52-57.

Room, R. (2005). The wheel of fortune: cycles and reactions in gambling policies. *Addiction, 100*(9), 1226-1227; discussion 1235. doi:10.1111/j.1360-0443.2005.01201.x

Room, R., Turner, N. E., & Ialomiteanu, A. (1999). Community effects of the opening of the Niagara casino. *Addiction, 94*(10), 1449-1466.

Rush, B., Veldhuizen, S., & Adlaf, E. (2007). Mapping the prevalence of problem gambling and its association with treatment accessibility and proximity to gambling venues. *Journal of Gambling Issues*(20), 193-215.

SEIGMA Research Team. (2018). *Social and economic impacts of expanded gambling in Massachusetts: 2018*. Retrieved from Amherst, MA:

Sevigny, S., Ladouceur, R., Jacques, C., & Cantinotti, M. (2008). Links between casino proximity and gambling participation, expenditure, and pathology. *Psychol Addict Behav, 22*(2), 295-301. doi:10.1037/0893-164X.22.2.295

Shaffer, H. J. (2005). From disabling to enabling the public interest: natural transitions from gambling exposure to adaptation and self-regulation. *Addiction, 100*(9), 1227-1230; discussion 1235. doi:10.1111/j.1360-0443.2005.01200.x

Shaffer, H. J., Hall, M. N., & Vander Bilt, J. (1997). *Estimating the Prevalence of Disordered Gambling Behavior in the United States and Canada: A Meta-analysis*. Retrieved from Boston:

Shaffer, H. J., & Simoneau, G. (2001). Reducing resistance and denial by exercising ambivalence during the treatment of addiction. *Journal of Substance Abuse Treatment, 20*(1), 99-105.

Storer, J., Abbott, M., & Stubbs, J. (2009). Access or adaptation? A meta-analysis of surveys of problem gambling prevalence in Australia and New Zealand with respect to concentration of electronic gaming machines. *International Gambling Studies, 9*(3), 225-244.

Toce-Gerstein, M., Gerstein, D. R., & Volberg, R. A. (2003). A hierarchy of gambling disorders in the general population. *Addiction, 98*, 1661-1672.

Volberg, R. A. (2004). Fifteen years of problem gambling research: What do we know? Where do we go? . *Journal of Gambling Issues, 10*.
Volberg, R. A., Gerstein, D. R., Christiansen, E. M., & Baldridge, J. (2001). Assessing self-reported expenditures on gambling. *Managerial and Decision Economics, 22*(1-3), 77-96.
Volberg, R. A., & Williams, R. J. (2014). Epidemiology: An international perspective. In D. C. S. Richard, A. Blaszczynski, & L. Nower (Eds.), *The Wiley-Blackwell handbook of disordered gambling* (pp. 26-48). Malden, MA: John Wiley & Sons.
Volberg, R. A., Williams, R. J., Stanek, E., Houpt, K. A., Zorn, M., & Rodriguez-Monguio, R. (2017). *Gambling and problem gambling in Massachusetts: Results of a baseline population survey*. Retrieved from Amherst, MA:
Volberg, R. A., & Wray, M. (2013). Beyond social control: Prevalence research and the uses of expert knowledge. *Addiction Research and Theory, 21*(1), 15-16.
Welte, J. W., Barnes, G. M., Tidwell, M. O., Hoffman, J. H., & Wieczorek, W. F. (2016). The Relationship Between Distance from Gambling Venues and Gambling Participation and Problem Gambling Among U.S. Adults. *J Gambl Stud, 32*(4), 1055-1063. doi:10.1007/s10899-015-9583-5
Welte, J. W., Wieczorek, W. F., Barnes, G. M., Tidwell, M. C., & Hoffman, J. H. (2004). The relationship of ecological and geographic factors to gambling behavior and pathology. *J Gambl Stud, 20*(4), 405-423. doi:10.1007/s10899-004-4582-y
Williams, R. J., Belanger, Y. D., & Arthur, J. N. (2011). *Gambling in Alberta: History, current status, and socioeconomic impacts*. Retrieved from Edmonton: https://dspace.ucalgary.ca/bitstream/1880/48495/1/SEIGA_Final_Report_April_2_2011.pdf
Williams, R. J., Hann, R., Schopflocher, D., West, B., McLaughlin, P., White, N., . . . Flexhaug, T. (2015). *Quinte longitudinal study of gambling and problem gambling*. Retrieved from Guelph: http://www.opgrc.org/sites/default/files/documents/QLS-OPGRC-2015.pdf
Williams, R. J., Rehm, J., & Stevens, R. M. G. (2011). *The social and economic impacts of gambling*. Retrieved from Winnipeg: http://hdl.handle.net/10133/1286
Williams, R. J., & Volberg, R. A. (2009). Impact of survey description, administration format, and exclusionary criteria on population prevalence rates of problem gambling. *International Gambling Studies, 9*(2), 101-117.
Williams, R. J., & Volberg, R. A. (2010). *Best practices in the population assessment of problem gambling*. Retrieved from Guelph: http://www.gamblingresearch.org/content/research.php?appid=2500
Williams, R. J., & Volberg, R. A. (2014). The classification accuracy of four problem gambling assessment instruments in population research. *International Gambling Studies, 14*(1), 15-28. doi:10.1080/14459795.2013.839731
Williams, R. J., Volberg, R. A., & Stevens, R. M. G. (2011). *The population prevalence of problem gambling: Methodological influences, standardized rates, jurisdictional differences, and worldwide trends*. Retrieved from Guelph:
Williams, R. J., Volberg, R. A., & Stevens, R. M. G. (2012). *The population prevalence of problem gambling: Methodological influences, standardized rates, jurisdictional differences, and worldwide trends*. Retrieved from Guelph: https://www.uleth.ca/dspace/handle/10133/3068
Williams, R. J., Volberg, R.A., & Stevens, R.M.G. (2012). *Population Assessment of Problem Gambling: Methodological Influences, Standardized Rates, Jurisdictional Differences, and Worldwide Trends.* Retrieved from
Williams, R. J., & Wood, R. T. (2007). The proportion of Ontario gambling revenue derived from problem gamblers. *Canadian Public Policy, 33*(3), 367-388.
Williams, R. J., Zorn, M., Stanek, E. J., Evans, V., & Volberg, R. A. (2020). *MA Gambling Impact Cohort (MAGIC): Transitions across four waves*. Retrieved from Amherst, MA:

Wood, R. T., & Williams, R. J. (2007). How much money do you spend on gambling? The comparative validity of question wordings used to assess gambling expenditure. *International Journal of Social Research Methodology, 10*(1), 63-77.

Young, M. (2013). Statistics, scapegoats and social control: A critique of pathological gambling prevalence research. *Addiction Research and Theory, 21*(1), 1-11.

Appendix A: Weighting Calculations

Appendix A1: Weighting Procedures for Springfield Baseline Target Population Survey

Introduction

The Springfield target survey selected a simple random sample of addresses from a sampling frame of address defined by Springfield and the surrounding communities. A simple random sample of addresses was selected, with the sample size of n=4,685 set to achieve a minimum of 1000 completed surveys. Weights are assigned to survey respondents that make the weighted responses representative of the Springfield target community population. A total of 1,131 surveys were completed between 2/23/2015 and 6/25/2015.

There are five steps in the weighing. First, a weight is assigned to directly account for the sampling fraction of addresses. Next, using information on the sample addresses, adjustments were made for unknown eligibility of the addresses. The third step in the weighting is to adjust for non-response. Household size is accounted for next. Finally, raking is used to align the weights with the distribution of four demographic variables (age, gender, race, and education) to the target population based on the 2015 American Community Survey PUMS data. The first three steps of the weighting were initially conducted by NORC, and revised at UMASS. The last two steps were conducted by the SEIGMA project at UMASS.

The weighting scheme involves the following steps:

1. Base sampling weight for sample of addresses;
2. Adjustment for unknown eligibility of addresses;
3. Adjustment for non-response to the questionnaire;
4. Adjustment for household size;
5. Raking of weights to age, gender, race, and education based on the 2015 PUMS.

Each individual weighting step is discussed in detail below.

Step 1. Base sampling weight

The base weight reflects the probability of an address being selected from the population frame. A total of 4,685 addresses were selected via simple random sampling from the 168,978 addresses in the frame defined by NORC, resulting in a probability of selection of $\pi = \frac{4,685}{168,978} = 0.0277$. The first weight is equal to the inverse of the probability of selection:

$$W_1 = \frac{1}{\pi} = 36.0679 \ .$$

Step 2. Adjustment for unknown eligibility

The first adjustment to the weights consists of an adjustment to account for those cases that were unable to be contacted or sufficiently screened, and thus have an unknown eligibility status. In this adjustment, the weight for cases of known eligibility status is adjusted to account for cases with unknown eligibility status. This assumes that the unknown cases would be eligible in the same proportion as the cases of known status, taking into account various other factors.

Among the sample addresses, a total of 1,192 were contacted/sufficiently screened so that they could be assessed as to the eligibility of the address. Thus, the unknown eligibility rate,

$$\pi_2 = \frac{Known\_Eligiblity}{(Known\_Eligiblity) + (Unknown\_Eligibility)}$$

is estimated by $\hat{\pi}_2 = \frac{1,192}{4,685} = 0.2544$.

Some additional information was available from the address frame so that it was possible to examine differences in the eligibility rate between addresses with different characteristics. Characteristics examined included type of record (highrise, street), type of address (single family dwelling, multi-family dwelling), language (address in English speaking area, address in Spanish speaking area), and delivery type (Curb line, Cluster Box Unit, Central, Other). An address was deemed "Spanish language area" if the area (specifically Census defined block group) had 25% or more of the households where Spanish was spoken.

Using eligibility (known) as an outcome, a stepwise logistic regression analysis was conducted to identify variables related to the screening rate. A p-value less than 0.15 was required for a variables to enter into the model, and a p-value less than 0.10 was required for a variable to stay in the model. Two variables entered the model, given in order by language and address type. The screening rates by a cross-classification of the address characteristics are given in Table A1.

Table A1. Screening Rates by Type of Address, Dwelling unit, and Language

Address Type	English		Spanish		All	
	N	Percent	N	Percent	N	Percent
Single Family	2,389	32.0%	886	19.1%	3,275	28.5%
Multiple Family	560	24.3%	850	14.5%	1,410	18.4%
All	2,949	30.5%	1,736	16.8%	4,685	25.4%

Table A2 summarizes the screening rates (for eligibility) by address groups. These rates are used to adjust weights for screening.

Table A2. Screening Rates by Address Groups

Address Groups	Eligibility Screened		
	Sample n_{2k}	Screened m_{2k}	Percent
Single Family - English	2,389	764	32.0%
Multiple Family - Spanish	886	169	19.1%
Multiple - Other - English	560	136	24.3%
Multiple - Other - Spanish	850	123	14.5%
All	4,685	1,192	25.4%

We estimate the screening rate for each group, $k=1,\ldots,K=4$, as $\hat{\pi}_{2k}=\frac{m_{2k}}{n_{2k}}$, where m_{2k} is the number of screened sampled addresses (where eligibility was known) in group k and n_{2k} is the number of sampled addresses in group k. We use the estimated screening rate to define the weight adjusted for screening, given by

$$W_{2k}=\frac{W_1}{\hat{\pi}_{2k}} .$$

This weight is assigned to all 1,192 screened addresses where eligibility is known, and is set to 'missing' for addresses with unknown eligibility status.

Step 3 Adjustment for Interview Completion Rate

The target respondent from each screened address was attempted to be interviewed, and the interview was classified as complete or not complete. A total of 1,131 interviews were complete, yielding a completion rate of $\hat{\pi}_3=\frac{1{,}131}{1{,}192}=0.9488$.

It was possible to examine differences in the completion rate using variables for type of record (highrise, street), type of address (single family dwelling, multi-family dwelling), language (address in English speaking area, address in Spanish speaking area), delivery type (Curb line, Cluster Box Unit, Central, Other), and attempted mode of interview (Web, Web/SAQ, Web/SAQ/Phone). Using whether or not the interview was complete as an outcome, a stepwise logistic regression analysis was conducted to identify variables related to the completion rate. A p-value less than 0.15 was required for a variables to enter into the model, and a p-value less than 0.10 was required for a variable to stay in the model. Two variables entered the model, given in order by the mode of interview, and language. The completion rates by a cross-classification of these variables is given in Table 3. We note that the total in Table A3 is 1,191, since the mode of the survey was missing for one screened sample address.

Table A3. Completion Rates by Mode of Interview Attempted and Language

Mode	English		Spanish		All	
	N	Percent	N	Percent	N	Percent
Web	397	100%	103	99.0%	500	99.8%
Web/SAQ	384	97.1%	136	94.9%	520	96.5%
Web/SAQ/Phone	118	81.4%	53	64.2%	171	76.0%
All	899	96.3%	292	90.8%	1,191	95.0%

We assigned an interview mode of Web only to this sample address resulting in the completion rates as given in Table A4.

Table A4. Completion Rates by Mode of Interview Attempted and Language

Mode	English		Spanish		All	
	N	Percent	N	Percent	N	Percent
Web	398	99.7%	103	99.0%	501	99.6%
Web/SAQ	384	97.1%	136	94.9%	520	96.5%
Web/SAQ/Phone	118	81.4%	53	64.2%	171	76.0%
All	900	96.2%	292	90.8%	1,192	94.9%

A total of 1,131 of the screened addresses had interviews that were completed, for a completion rate of 94.9%. We totaled the screening weights for addresses with completed interviews, and where interviews were not completed by the mode of interview and language. These totals are given in Table A5.

Table A5. Total Screening Weights by Completion Status, Mode of interview, and Language

		Not started		Complete		All	
		# Subjects	Total SP1WT2	# Subjects m^*_{3h}	Total SP1WT2	# Subjects n^*_{3h}	Total SP1WT2
Web	English	1	149	397	47,133	398	47,282
	Spanish	1	249	102	22,175	103	22,424
Web/SAQ	English	11	1,241	373	44,033	384	45,274
	Spanish	7	1,504	129	27,401	136	28,905
Web/SAQ/Phone	English	22	2,588	96	11,220	118	13,809
	Spanish	19	4,074	34	7,211	53	11,285
All	-	61	9,805	1,131	159,173	1,192	168,978

We estimate the completion rates for each group, $h=1,\ldots,H=6$, as $\hat{\pi}_{3h}=\frac{m^*_{3h}}{n^*_{3h}}$, where $m^*_{3h}=\sum_{j=1}^{m_{3h}} W_{2hj}$ is the sum of the weights W_{2hj} for the $j=1,\ldots,m_{3h}$ respondents who completed the survey, $n^*_{3h}=\sum_{j=1}^{n_{3h}} W_{2hj}$ is the sum of the weights W_{2hj} for the $j=1,\ldots,n_{3h}$ respondents who completed the screening. We use the estimated completion rate to define the weight adjusted for screening, and non-response given by

$$W_{3hk}=\frac{W_{2k}}{\hat{\pi}_{3h}}$$

for $k=1,\ldots,n_3=1{,}131$ completed survey respondents. This weight is assigned to all 1,131 completed survey respondents, and is set to 'missing' for addresses where the survey was not complete.

Step 4 Adjustment for Household Size

The fourth adjustment in the weights is for household size. The number of persons 18 years or older living the household was recorded for most (1,100 or 97.3%) of the completed interviews. The distribution of household size for completed respondents is given in Table A6.

Table A6. Number of 18+ Adults in Household

# 18+ Adults in Household	Frequency	Percent	Cumulative Frequency	Cumulative Percent
-	31	-	-	-
1	369	33.55%	369	33.55%
2	533	48.45%	902	82.00%
3	127	11.55%	1029	93.55%
4	55	5.00%	1084	98.55%
5	10	0.91%	1094	99.45%
6	4	0.36%	1098	99.82%
8	1	0.09%	1099	99.91%
9	1	0.09%	1100	100.00%

Let the household size (based on 18+ age household members) for respondent j be represented by p_j. The total number of addresses in the Springfield frame is estimated for the 1,131 respondents as $\sum_{j=1}^{1131} W_{3j} = 168,978$. This total can be divided into addresses where the number of 18+ persons in the household is known, $\sum_{j=1}^{1100} W_{3j} = 164,862$, and addresses where the number of 18+ persons is unknown, $\sum_{j^*=1}^{31} W_{3j^*} = 4,115$. Using the household size when it is known, the total number of persons in the 164,862 households is estimated by $\sum_{j=1}^{1100} W_{3j} p_j = 312,248$, resulting in an average household size of $\frac{312,248}{164,862} = 1.894$, which we round to 2. We assign this household size to the 31 respondents where household size is unknown. We further truncate the household size to a maximum of 4 in an effort to limit the variability of the survey weights. The weight adjusted for household size is given by

$$W_{4,j}^* = p_j W_{3,j}.$$

The total weight corresponds to the estimated number of 18+ persons in the target area, given by $\sum_{j=1}^{1131} W_{4j}^* = 315,932$.

In the Springfield Target area (defined by PUMA codes 1900,1901,1902, and 1600), the total number of persons age 18+ based on the 2015 PUMS data is 454,772 (see ged18p002.sas). We adjust the total number of 18+ persons to match the PUMS total to define the household size adjusted weight given by

$$W_{4,j} = \frac{454,772}{315,932} W_{4,j}^*.$$

Step 5 Aligning to 2015 Population Via Raking

We determine a set of raking variables after review of data from the 2015 ACS PUMS data set, and review of the survey respondent data. For the PUMS data, the review considered the adequacy of the PUMS data for raking cells by examining the number of respondents in a cell, and the coefficient of variation of the mean (CVM) weight (statistical weight assigned in the ACS) to a cell. When the number of respondents is small (less than 20), and/or the CVM is large (greater than 10%), some cells were combined so as to produce stable estimates of the total persons in the cell. When the CVM is less than 10%, we consider the uncertainty in estimating the total number of adults in the cell to be acceptably small. In this situation, we consider the assumption of ignoring the uncertainty in the population total weight estimates for each cell to be justified.

In addition to the investigation of the adequacy of data in cells from the PUMS, we examined the number of respondents in each cell in the target survey. Cells where there were fewer than 10 respondents were identified, and considered for possible collapsing.

Primary Variables for Raking Weights for Targeted Areas using the 2015 PUMS Population Data

We consider four variables: age (18-34, 35-49, 50-64, 65+), gender (male, female), race/ethnicity (Hispanic, Black (only), Asian (only), White/other), and education (high school or less, some college/college graduate, some postgraduate education). We consider these variables individually and pair-wise. For some variables, we also consider a collapsed set of categories (i.e., for race/ethnicity: Hispanic/ Asian and Black/Asian, and for education: high school or less, or more than high school).

We determined raking variables via a preliminary analysis of the 2015 one-year American Community Survey Public Use Microdata Sample (PUMS) files. In an ideal setting, reliable PUMS data for population totals would be available for a full cross-classification of weighting variables. In practice, estimates of the population based on the PUMS data are based on an approximate 1% sample of the MA population and the PUMS data themselves are weighted to estimate the number of individuals in each post-stratum. We evaluate the reliability of the PUMS estimates of totals in cells formed by individual variables and a cross-classification of pairs of the variables. Reliability was evaluated by: (1) the number of PUMS respondents for each possible raking cell and (2) calculating the CVM in the estimated total PUMS weight. The CVM is the estimated standard error of the mean, divided by the mean, and expressed as a percent.

Table A7 provides the CVM and the number of respondents in the 2015 PUMS for the Springfield target area for each of the four variables used in weighting in the BTPS. The CVM was less than 10% for all cells, and in all cells, there are greater than 20 respondents.

Table A7. CVM and N by Age, Gender, Race, and Education for Springfield Using 2015 PUMS Data

PUMS 2015	Age				Gender		Race			
	18-34	35-49	50-64	65+	Male	Female	Hispanic	White/Other	Black	Asian
CVM (Wt)	2.3%	2.2%	1.8%	1.9%	1.5%	1.5%	2.9%	1,2%	5.0%	7.8%
n	1,263	1,022	1,315	1,103	2,151	2,552	612	3,720	261	110

PUMS 2015	Education		
	<= High School	College	Graduate School
CVM (Wt)	1.7%	1.5%	2.7%
n	1,851	2,254	598

Next, we consider pair-wise cross-classifications of the weighting variables by region (Tables A8-A10). Once again, we highlight cells (in red) where the number of respondents is less than 20 and the CVM is greater than 10%.

Table A8. CVM and N by Two Variables (Age and one other) for Springfield Using 2015 PUMS Data

PUMS 2015			Age			
			18-34	35-49	50-64	65+
Gender	Male	CVM (Wt)	3.2%	3.1%	2.7%	2.7%
		n	583	484	622	462
	Female	CVM (Wt)	3.3%	3.0%	2.5%	2.7%
		n	680	538	693	641
Education	<= High School	CVM (Wt)	3.7%	3.9%	3.1%	2.6%
		n	434	343	515	559
	College	CVM (Wt)	3.1%	3.0%	2.6%	3.4%
		n	754	503	619	378
	Graduate School	CVM (Wt)	8.9%	4.3%	4.5%	5.1%
		n	75	176	181	166
Race	Hispanic	CVM (Wt)	4.6%	5.2%	5.6%	8.8%
		n	267	164	120	61
	White/Other	CVM (Wt)	2.7%	2.4%	2.0%	2.0%
		n	872	763	1,101	984
	Black	CVM (Wt)	10.7%	8.9%	7.9%	9.2%
		n	77	63	71	50
	Asian	CVM (Wt)	14.1%	13.8%	12.8%	27.5%
		n	47	32	23	8

Table A9. CVM and N by Two Variables (Race and one other) for Springfield Using 2015 PUMS Data

PUMS 2015			Race			
			Hispanic	White/Other	Black	Asian
Gender	Male	CVM (Wt)	4.2%	1.7%	7.5%	12.0%
		n	275	1,705	127	44
	Female	CVM (Wt)	3.9%	1.6%	6.5%	10.4%
		n	337	2,015	134	66
Education	<= High School	CVM (Wt)	3.5%	2.0%	7.6%	10.5%
		n	394	1,297	117	43
	College	CVM (Wt)	5.2%	1.6%	7.2%	15.5%
		n	202	1,890	119	43
	Graduate School	CVM (Wt)	16.0%	2.8%	14.5%	12.7%
		n	16	533	25	24

Table A10. CVM and N by Two Variables (Gender x Education) for Springfield Using 2015 PUMS Data

PUMS 2015			Education		
			<= High School	College	Graduate School
Gender	Male	CVM (Wt)	2.4%	2.3%	4.0%
		n	947	954	250
	Female	CVM (Wt)	2.5%	2.1%	3.6%
		n	904	1,300	348

For Springfield, the results in Tables A8-A10 indicate that using four categories for race will result in a large CVM for Asians when cross-classified by age (Table A8), gender (Table A9), and education (Table A10). We examined the distribution of the variables age, gender, and education with race, in an effort to identify a racial group most similar to Asians. The results are given (using the weighted PUMS data) in Table A11.

Table A11. Estimated Totals by Race for Age, Gender, and Education Using 2015 PUMS Data

PUMS 2015			Race				
			Hispanic	White/Other	Black	Asian	Total
Age	18-34	Frequency	34,514	86,533	9,406	4,362	134,815
		Column %	44.84%	25.53%	33.54%	40.62%	
	35-49	Frequency	22,705	73,020	7,667	3,264	106,656
		Column %	29.50%	21.54%	27.34%	30.40%	
	50-64	Frequency	13,276	99,478	6,888	2,557	122,199
		Column %	17.25%	29.34%	24.56%	23.81%	
	65+	Frequency	6,481	79,981	4,085	555	91,102
		Column %	8.42%	23.59%	14.57%	5.17%	
Total			76,976	399,012	28,046	10,738	454,772
Gender	Male	Frequency	35,088	159,849	14,192	4,216	213,345
		Column %	45.58%	47.15%	50.60%	39.26%	
	Female	Frequency	41,888	179,163	13,854	6,522	241,427
		Column %	54.42%	52.85%	49.40%	60.74%	
Total			76,976	339,012	28,046	10,738	454,772
Education	<= High School	Frequency	50,220	124,844	12,374	4785	192,223
		Column %	65.24%	36.83%	44.12%	44.56%	
	College	Frequency	24,944	167,819	13,672	4,050	210,485
		Column %	32.40%	49.50%	48.75%	37.72%	
	Graduate School	Frequency	1,812	46,349	2,000	1,903	52,064
		Column %	2.35%	13.67%	7.13%	17.72%	
Total			76,976	399,012	28,046	10,738	454,772

The results in Table A11 suggest similar distributions of age and gender for Hispanics and Asians, but more similar distributions of education with blacks, particularly after collapsing the categories of college and graduate education. We created additional variables to reflect these combinations of race categories. One variable, P_RACEV1, combines Hispanics and Asians into a single group. A second variable, P_RACEV2, combines Blacks and Asians into a single group. A third variable combines college and graduate education into a single group (P_EDUV1). Tables A12-A14 re-evaluates the CVM and number of subjects using these variables when cross-classified.

Table A12. CVM and N by Two Variables (Age and Collapsed Race) for Springfield Using 2015 PUMS Data

PUMS 2015			Age			
			18-34	35-49	50-64	65+
Race	Hispanic/Asian	CVM (Wt)	4.4%	4.9%	5.1%	8.5%
		n	314	196	143	69
	White/Other	CVM (Wt)	2.7%	2.4%	2.0%	2.0%
		n	872	763	1,101	984
	Black	CVM (Wt)	10.7%	8.9%	7.9%	9.2%
		n	77	63	71	50

Table A13. CVM and N by Two Variables (Gender and Collapsed Race) for Springfield Using 2015 PUMS Data

PUMS 2015			Gender	
			Male	Female
Race	Hispanic/Asian	CVM (Wt)	4.0%	3.7%
		n	319	403
	White/Other	CVM (Wt)	1.7%	1.6%
		n	1,705	2,015
	Black	CVM (Wt)	7.5%	6.5%
		n	127	134

Table A14. CVM and N by Two Variables (Collapsed Education and Collapsed Race) for Springfield Using 2015 PUMS Data

PUMS 2015			Education	
			<= High School	> High School
Race	Hispanic/Asian	CVM (Wt)	3.5%	5.0%
		n	394	218
	White/Other	CVM (Wt)	2.0%	1.4%
		n	1,297	2,423
	Black	CVM (Wt)	6.2%	5.7%
		n	160	211

For Springfield, after collapsing race, and education as indicated in Tables A12-A14 above, all cells contain at least 20 respondents and have a CVM less than 10% (with the exception of 18-34 year old Blacks (Table A12) where the CVM is 10.7%).

Respondents in the Springfield Target Survey

In addition to having stable estimates of the population total in each raking cell, we examine the number of respondents in the Target survey by possible raking variables if there are few respondents in a region in a particular category, the weight assigned to the subjects may be difficult to match to the population weights. When there are no respondents in a cell, the cell needs to be collapsed to match population weights. Table A15 on the following page considers each raking variable individually. Note that there are more than 10 respondents in each non-missing cell.

Table A15. Number of Respondents by Demographics for Springfield Targeted Survey 2015

Respondents 2015 Survey		n
Age	18-34	175
	35-49	229
	50-64	322
	65+	351
	Missing	54
Gender	Male	411
	Female	715
	Missing	5
Education	<= High School	259
	College	590
	Graduate School	264
	Missing	18
Race	Hispanic	109
	White/Other	895
	Black	57
	Asian	20
	Missing	50

The results in Table A15 indicate that there are greater than 10 respondents in each demographic cell in the target survey (except for cells where the variable was missing). Nevertheless, there were relatively few respondents for categories of race that were Asian. For this reason, we collapsed race in a similar manner as was done for the 2015 PUMS data when cross-classifying variables (Table A16-A19).

Table A16. Number of Respondents by Demographics for Springfield Targeted Survey 2015

Respondents 2015 Survey		Age				
		18-34	35-49	50-64	65+	Missing
		n	n	n	n	n
Gender	Male	53	75	129	136	18
	Female	122	154	193	215	31
	Missing	-	-	-	5	-
Race	Hispanic/Asian	47	24	33	18	7
	White/Other	107	187	263	309	29
	Black	12	14	16	14	1
	Missing	9	4	10	10	17
Education	<= High School	44	28	67	114	6
	College	101	133	179	155	22
	Graduate School	29	67	72	76	20
	Missing	1	1	4	6	6

Table A17. Number of Respondents by Demographics for Springfield Targeted Survey 2015

Respondents 2015 Survey		Education			
		<= High School	College	Graduate School	Missing
		n	n	n	n
Gender	Male	90	215	100	6
	Female	169	375	163	8
	Missing	-	-	1	4

Table A18. Number of Respondents by Demographics for Springfield Targeted Survey 2015

Respondents 2015 Survey		Education		
		<= High School	> High School	Missing
		n	n	n
Race	Hispanic	43	64	2
	White/Other	186	707	2
	Black/Asian	21	56	-
	Missing	9	27	14

Table A19. Number of Respondents by Demographics for Springfield Targeted Survey 2015

Respondents 2015 Survey		Race			
		Hispanic/Asian	White/Other	Black	Missing
		n	n	n	n
Gender	Male	36	340	16	19
	Female	93	554	41	27
	Missing	-	1	-	4

The results in Tables A16-A19 indicate that there are greater than 10 respondents in each demographic cell in the two-way classifications when variables are collapsed similar to those in the PUMS data.

Raking Weights for the Springfield Target Survey using the 2015 ACS Population Data

We adjusted weights assigned to subjects to more closely align with the distribution of 18+ year old persons in MA for Springfield using a raking procedure. Based on data from the 2015 American Community Survey Public Use Microdata Sample (PUMS) files, and the distribution of respondents to the Springfield Target survey, adjustments were made for age (18-34, 35-49, 50-64, 65+), gender (male, female), race/ethnicity (Hispanic, Black (only), Asian (only), White and other), education (high school or less, some college/college graduate, some post graduate education), and cross-classifications of variables as indicated in Table A19. For cross-classifications that included race and age or gender, categories for Hispanics and Asians are combined. For cross-classifications that included race and education, categories for Blacks and Asians are combined, and categories for college and graduate education are combined. The weights are developed are based on raking using a similar procedure as was used for developing raked weights in the Plainville Baseline Target Survey. The raking variables used are V1-V7 as indicated in Table A20.

Table A20. Variables Using in Raking for Springfield Target Survey

Raking Variable	V1	V2	V3	V4	V5	V6	V7
Variables (# of Cells)	Race (5)	Age x Gender (15)	Age x RaceV1 (20)	Age x Education (20)	Gender x RaceV1 (12)	Gender x Education (12)	RaceV2 x EducV1 (12)
Categories 1st Variable	Hispanic	18-34	18-34	18-34	Male	Male	Hispanic
	White/Other	35-49	35-49	35-49	Female	Female	White/Other
	Black	50-64	50-64	50-64	Missing	Missing	Black/Asian
	Asian	65+	65+	65+			Missing
	Missing	Missing	Missing	Missing			
Categories 2nd Variable		Male	White/Other	HS or less	White/Other	HS or less	HS or less
		Female	Black	College	Black	College	Some College
		Missing	Hispanic/Asian	Graduate	Hispanic/Asian	Graduate	Missing
			Missing	Missing	Missing	Missing	

We first account for missing values for race, age x gender, age x collapsed race, age x education, gender x collapsed race, gender x education, and collapsed race x collapsed education by assigning weights to respondents with some missing values equal to the weights in the Springfield target survey. This process is explained in more detail below.

Accounting for Missing Data among Respondents for Raking Variables

Several steps are taken to develop raked weights. First, a weight is assigned to respondents who are missing response for each of the raked variables. Next, the total weight assigned to categories of a variable for respondents is matched to the total weight in the 2015 PUMS data.

We begin assuming that weights have been developed for respondents in the Springfield target survey such that the total weight for respondents matches the total adult population defined by the PUMA codes for the Springfield area in 2015 (i.e. 454,772). This weight is sp1wt4. Seven variables are considered for raking, with some variables defined by a cross-classification of two other variables. For each variable, we add a missing value category. Then, using the augmented categories, we create a variable that uniquely defines the age (5) x gender (3) x Race (5) x Education (4) x Collapsed Education (3) x Collapsed RaceV1 (4) x Collapsed RaceV2 (4) levels. This variable has 14,400 categories. A similar variable is created using the 2015 PUMS data for the area. Using these categories, the total weight (sp1wt4) is summed for each category using the Springfield data, and using the PUMS data (using the pwgpt variable). These two sets of counts (each totaling the 2015 population total) are the input for the raking.

Raking is an iterative process, with 1 iteration corresponding to an attempt to match marginal for each of the tables. For Springfield, we begin with the 7 raking variables:

Table A21. List of Raking Variables and Number of Categories (Including Missing Categories) for Raking in Springfield

Variables	Categories	Number of Categories
V1	Race/Ethnicity	5
V2	Age x Gender	15
V3	Age x Collapsed Race	15
V4	Age x Education	20
V5	Gender x Collapsed Race	9
V6	Gender x Education	12
V7	Collapsed Race x Collapsed Education	12

We use the indices $i=1,...,5$ for categories of age (including 'missing age' as $i=5$), $j=1,...,3$ for categories of gender (including 'missing gender' as $j=3$), $k=1,...,5$ for categories of race/ethnicity (including 'missing race/ethnicity' as $k=5$), $l=1,...,4$ for categories of education (including 'missing education' as $l=4$), $m=1,...,3$ for categories of collapsed education (including 'missing collapsed education' as $m=3$), $n=1,...,4$ for categories of collapsed racev1 (including 'missing collapsed race' as $n=4$), and $p=1,...,4$ for categories of collapsed racev2 (including 'missing collapsed race' as $p=4$). We illustrate the process of raking weights using the first raking variable, V1, corresponding to race/ethnicity, and then describe the overall raking process. Raking was accomplished using a SAS program written for this purpose. The first step was to evaluate the total weight (NSP1WT4) in each of the 5 cells for the sample. Let us refer to these weights by x_k for $k=1,...,5$ (corresponding to race, including ‘5’ as a missing age category), so that the total sample weight is given by $x_+ = \sum_{k=1}^{5} x_k$, where $x_k = \sum_{i=1}^{5}\sum_{j=1}^{3}\sum_{l=1}^{4}\sum_{m=1}^{3}\sum_{n=1}^{4}\sum_{p=1}^{4} x_{ijklmnp}$. The population weights, p_k , are based on the 2015 PUMS data. Among the population data, there are no missing values. Using the categories of race, the total population is the sum over 4 race cells, $p_+ = \sum_{k=1}^{4} p_k$, where $p_k = \sum_{i=1}^{5}\sum_{j=1}^{3}\sum_{l=1}^{4}\sum_{m=1}^{3}\sum_{n=1}^{4}\sum_{p=1}^{4} p_{ijklmnp}$. As a result, when raking by the variable V1, we first re-allocated PUMS data to form categories representing “missing age.” Table A22 illustrates these totals for Springfield prior to and after adjusting for missing race/ethnicity data.

Table A22. Initial Weights and Missing-Adjusted PUMS Weights for Race for Springfield (2015 PUMS) V1

Categories		Springfield Wt	PUMS Wt	Rev PUMS Wt
		Sum	Sum	Sum
Race	Hispanic	65,364	76,976	73,742
	White/Other	334,949	339,012	324,767
	Black	25,426	28,046	26,868
	Asian	9,924	10,738	10,287
	Missing	19,109	0	19,109
	All	454,772	454,772	454,772

The population weight for the missing data is calculated by assigning population weights to cells where race is missing proportional to the weight assigned these cells in the sample, $p_5^* = p_+\left(\frac{x_5}{x_+}\right)$. Since $p_+ = x_+$ (except for differences due to rounding) , $p_5^* = x_5 = 19,109$ (representing weights assigned to the population with missing race). Weights for other age categories are adjusted, to preserve the overall population weight, p_+ , such that $p_i^* = p_i\left(\frac{p_+ - p_5^*}{p_+}\right)$, for $i = 1,...,4$. For example, the adjusted 2015 PUMS weight for Hispanics in Springfield is given by $73,742 = 76,967\left(\frac{454,772 - 19,109}{454,772}\right)$. Similar results are given for each of the other raking variables (Tables A23-A28).

Table A23. Weights Accounting for Missing Values for Age x Gender for Springfield Follow-up (2015 PUMS) V2

Age Categories		Age					
		18-34	35-49	50-64	65+	Missing	All
		Total Wt Sum	Total Wt Sum	Total Wt Sum	Total Wt Sum	Total Wt Sum	Total Wt Sum
PUMS 2015	Male	63,656	48,291	56,490	37,346	6,572	212,355
	Female	65,152	53,574	60,242	49,592	11,747	240,307
	Missing	0	0	0	0	2,110	2,110
	All	128,808	101,865	116,732	86,938	20,430	454,772
Springfield	Male	27,993	29,105	55,459	47,834	6,572	166,963
	Female	56,345	66,013	80,675	70,918	11,747	285,699
	Missing	0	0	0	0	2,110	2,110
	All	84,338	95,118	136,135	118,752	20,430	454,772

Table A24. Weights Accounting for Missing Values in V3: Age x RaceV1 for Springfield Follow-up (2015 PUMS)

Race Categories		Age					
		18-34	35-49	50-64	65+	Missing	All
		Total Wt Sum	Total Wt Sum	Total Wt Sum	Total Wt Sum	Total Wt Sum	Total Wt Sum
PUMS 2015	Hispanic/Asian	35,483	24,180	14,495	6,427	3,540	84,124
	White/Other	79,902	68,783	92,133	73,910	9,941	324,670
	Black	8,859	7,366	6,507	3,850	287	26,869
	Missing	4,394	1,496	3,647	2,909	6,662	19,109
	All	128,638	101,826	116,782	87,096	20,430	454,772
Springfield	Hispanic/Asian	27,523	13,874	20,845	9,507	3,540	75,288
	White/Other	47,164	74,134	105,109	98,602	9,941	334,949
	Black	5,257	5,614	6,534	7,734	287	25,426
	Missing	4,394	1,496	3.647	2,909	6,662	19,109
	All	84,338	95,118	136,135	118,752	20,430	454,772

Table A25. Weights Accounting for Missing Values in V4: Age x Edu for Springfield Follow-up (2015 PUMS)

Education Categories		Age					
		18-34	35-49	50-64	65+	Missing	All
		Total Wt Sum	Total Wt Sum	Total Wt Sum	Total Wt Sum	Total Wt Sum	Total Wt Sum
PUMS 2015	<= High School	53,074	39,341	48,926	46,248	1,698	189,286
	College	68,580	48,775	52,984	27,885	9,624	207,847
	Graduate School	7,272	12,987	13,451	10,728	6,888	51,326
	Missing	326	326	1,140	2,302	2,219	6,313
	All	129,252	101,428	116,500	87,163	20,430	454,772
Springfield	<= High School	23,683	14,233	33,492	39,942	1,698	113,047
	College	49,882	55,823	71,662	52,437	9,624	239,428
	Graduate School	10,448	10,448	24,736	24,070	6,888	95,984
	Missing	326	326	1,140	2,302	2,219	6,313
	All	84,338	95,118	136,135	118,752	20,430	454,772

Table A26. Weights Accounting for Missing Values in V5: Gender x RaceV1 for Springfield Follow-up (2015 PUMS)

Race Categories		Gender			
		Male	Female	Missing	All
		Total Wt Sum	Total Wt Sum	Total Wt Sum	Total Wt Sum
PUMS 2015	Hispanic/Asian	37,831	46,177	0	84,008
	White/Other	153,608	170,621	550	324,780
	Black	13,660	13,215	0	26,875
	Missing	7,255	10,294	1,560	19,109
	All	212,354	240,308	2,110	454,772
Springfield	Hispanic/Asian	20,541	54,747	0	75,288
	White/Other	132,697	201,702	550	334,949
	Black	6,471	18,955	0	25,426
	Missing	7,255	10,294	1,560	19,109
	All	166,963	285,699	2,110	454,772

Table A27. Weights Accounting for Missing Values in V6: Gender x Edu for Springfield Follow-up (2015 PUMS)

Education Categories		Gender			
		Male	Female	Missing	All
		Total Wt Sum	Total Wt Sum	Total Wt Sum	Total Wt Sum
PUMS 2015	<= High School	97,987	91,539	0	189,525
	College	91,373	116,205	0	207,578
	Graduate School	20,475	30,330	550	51,356
	Missing	2,557	2,197	1,560	6,313
	All	212,391	240,271	2,110	454,772
Springfield	<= High School	38,924	74,124	0	113,047
	College	87,534	151,894	0	239,428
	Graduate School	37,948	57,485	550	95,984
	Missing	2,557	2,197	1,560	6,313
	All	166,963	285,699	2,110	454,772

Table A28. Weights Accounting for Missing Values in V7: RaceV2 x EduV1 for Springfield Follow-up (2015 PUMS)

Education Categories		Race				
		Hispanic	White/Other	Black/Asian	Missing	All
		Total Wt Sum	Total Wt Sum	Total Wt Sum	Total Wt Sum	Total Wt Sum
PUMS 2015	<= High School	47,948	120,738	16,626	4,020	189,331
	> High School	25,046	203,074	20,543	10,465	259,128
	Missing	1,083	607	0	4,623	6,313
	All	74,077	324,418	37,169	19,109	454,772
Springfield	<= High School	28,268	71,222	9,537	4,020	113,047
	> High School	36,013	263,121	25,813	10,465	335,412
	Missing	1,083	607	0	4,623	6,313
	All	65,364	334,949	35,350	19,109	454,772

Matching Sample to Population Marginals for Steps with Raking Variables

The next step is to iteratively alter the respondent's weights so that the sample total weight for each raking table matches the population weight for the table. We illustrate this with an example. The totals for the Springfield Sample and the PUMS population differ for different categories of the raking variables. As an example from Table A28, the estimated number of Hispanics who had a High School or less education based on the Springfield sample is 28,268, while the estimated number from the Springfield PUMS is 47,948. The goal of raking is to alter the weights so that the Springfield sample cell totals are as similar as possible for the Springfield PUMS totals.

Description of a Step in the Raking

Raking is accomplished using a SAS program written for this purpose. The process proceeds in an iterative manner. Each iteration consists of a sequence of steps, where each step aligns the sample and population weighs for a raking variable. We describe the process for the first raking variable, V1, corresponding to race/ethnicity.

The total sample weight assigned to a cell for a raking variable is the sum of SP1WT4 assigned to respondents in that cell. We index categories for the 4 primary variables by $i=1,\dots,5$ for age, $j=1,\dots,3$ for gender, $k=1,\dots,5$ for race, $l=1,\dots,4$ for education, $m=1,\dots,3$ for collapsed education, $n=1,\dots,4$ for collapsed raceV1, and $p=1,\dots,4$ for collapsed raceV2. Respondents within a cell are indexed by $q=1,\dots,n_{ijklmnp}$. The total sample weight assigned to a cell for the first raking variable, V1 (i.e. race), is given by

$$x_k = \sum_{i=1}^{5}\sum_{j=1}^{3}\sum_{l=1}^{4}\sum_{m=1}^{3}\sum_{n=1}^{4}\sum_{p=1}^{4} x_{ijklmnp}$$
$$= \sum_{i=1}^{5}\sum_{j=1}^{3}\sum_{l=1}^{4}\sum_{m=1}^{3}\sum_{n=1}^{4}\sum_{p=1}^{4}\left(\sum_{q=1}^{n_{ijklmnp}} x_{ijklmnpq}\right),$$

where $\sum_{q=1}^{n_{ijklmnp}} x_{ijklmnpq}$, and $x_{ijklmnpq}$ is the value of SP1WT4 for respondent q. The first step in an iteration of raking aligns the sample marginal to the population marginal by forming the new weight for cells based on the full cross-classification of the raking variables, such that

$$x_{ijklmnp}^{(1)} = x_{ijklmnp}\left(\frac{p_k^*}{x_k}\right),$$

and p_k^* is the PUMS weight adjusted for missing data.

Using these weights, the total weight is evaluated for each cell for the raking variable, and for each cell corresponding to the remaining raking variables, such as V2 (corresponding to age x gender), etc, i.e. $x_{ij}^{(1)} = \sum_{k=1}^{5}\sum_{l=1}^{4}\sum_{m=1}^{3}\sum_{n=1}^{4}\sum_{p=1}^{4} x_{ijklmnp}^{(1)}$. Once again, using the population marginal weights, we align the sample marginal to the population marginal for V2, such that

$$x_{ijklmnp}^{(2)} = x_{ijklmnp}^{(1)}\left(\frac{p_{ij}^*}{x_{ij}^{(1)}}\right).$$

This process is continued for each of the raking variables, resulting in the marginal total weights in each cell after one iteration given by $r_{ijklmnp}^{1} = x_{ijklmnp}^{(7)}$.

Table A29 summarizes the sample and aligned population weights prior to a raking iteration, but after accounting for missing data. Table A30 illustrates the marginal weight totals for one variable, age x gender, on the 7 steps for the first three iterations of raking.

Table A29. Springfield Follow-up Targeted Survey: Sample and Aligned Population Weights Prior to Raking

V1	Race				
	Hispanic	White	Black	Asian	Missing
PUMS	73,742	3247,67	26,868	10,287	19,109
Sample	65,364	334,949	25,426	9,924	19,109

V2	Age/Gender														
	18-34			35-49			50-64			65+			Missing		
	Male	Female	Missing	Male	Female	Missing	Male	Female	Missing	Male	Female	Missing	Male	Female	Missing
PUMS	63,656	65,152	0	48,291	53,574	0	56,490	6,0242	0	37,346	49,592	0	6,572	11,747	2,110
Sample	27,993	56,345	0	29,105	66,013	0	56,490	80,675	0	47,834	70,918	0	6,572	11,747	2,110

V3	Age/Race											
	18-34				35-49				50-64			
	White	Black	Hispanic/Asian	Missing	White	Black	Hispanic/Asian	Missing	White	Black	Hispanic/Asian	Missing
PUMS	35,483	79,902	8,859	4,394	24,180	68,783	7,366	1,496	14,495	92,133	6,507	3,647
Sample	27,523	47,164	5,257	4,394	13,874	74,134	5,614	1,496	20,845	105,109	6,534	3,647

V3 (cont.)	Age/Race							
	65+				Missing			
	White	Black	Hispanic/Asian	Missing	White	Black	Hispanic/Asian	Missing
PUMS	6,427	73,910	3,850	2,909	3,540	9,941	287	6,662
Sample	9,507	98,602	7,734	2,909	3,540	9,941	287	6,662

V4	Age/Race											
	18-34				35-49				50-64			
	<= High School	College	Graduate School	Missing	<= High School	College	Graduate School	Missing	<= High School	College	Graduate School	Missing
PUMS	53,074	68,580	7,272	326	39,341	48,775	12,987	326	48,926	52,984	13,451	1,140
Sample	23,683	49,882	10,448	326	14,233	55,823	24,736	326	33,492	71,662	29,841	1,140

V4 (cont.)	Age/Race							
	65+				Missing			
	<= High School	College	Graduate School	Missing	<= High School	College	Graduate School	Missing
PUMS	46,248	27,885	10,728	2,302	1,698	9,624	6,888	2,219
Sample	39,942	52,437	24,070	2,302	1,698	9,624	6,888	2,219

V5	Gender/Race											
	Male				Female				Missing			
	White	Black	Hispanic/Asian	Missing	White	Black	Hispanic/Asian	Missing	White	Black	Hispanic/Asian	Missing
PUMS	37,831	153,608	13,660	7,255	46,177	170,621	13,215	10,294	0	550	0	1,560
Sample	20,541	132,697	6,471	7,255	54,747	201,702	18,955	10,294	0	550	0	1,560

V6	Gender/Education											
	Male				Female				Missing			
	<= High School	College	Graduate School	Missing	<= High School	College	Graduate School	Missing	<= High School	College	Graduate School	Missing
PUMS	97,987	91,373	20,475	2,557	91,539	116,205	30,330	2,197	0	0	550	1,560
Sample	38,924	87,534	37,948	2,557	74,124	151,894	57,485	2,197	0	0	550	1,560

V7	Race/Education											
	Hispanic			White/Other			Black/Asian			Missing		
	<= High School	> High School	Missing	<= High School	> High School	Missing	<= High School	> High School	Missing	<= High School	> High School	Missing
PUMS	47,948	25,046	1,083	120,738	203,074	607	16,626	20,543	0	4,020	10,465	4,623
Sample	28,268	36,013	1,083	71,222	263,121	607	9,537	25,813	0	4,020	10,465	4,623

	Max1	Max2	Max3	Max4	Max5	Max5	Max6	Max7
Max Percent Difference	124.4%	11.4%	56.0%	100.9%	124.4%	52.6%	89.5%	43.8%

Table A30. Springfield Follow-up Targeted Survey: Sample and Aligned Population Weights by Step and Iteration for V5: Age x Gender

Steps	Age by Gender V2														
	18-34			35-49			50-64			65+			Missing		
	Male	Female	Missing	Male	Female	Missing	Male	Female	Missing	Male	Female	Missing	Male	Female	Missing
PUMS	63,656	65,152	0	48,291	53,574	0	56,490	60,242	0	37,346	49,592	0	6,572	11,747	2,110
Sample	27,993	56,345	0	29,105	66,013	0	55,459	80,675	0	47,834	70,918	0	6,572	11,747	2,110
Iteration 1															
V1	28,262	58,104	0	28,653	66,031	0	55,086	80,841	0	47,028	70,288	0	6,499	11,886	2,093
V2	64,267	67,185	0	47,542	53,589	0	56,110	60,365	0	36,717	49,152	0	6,499	11,886	2,093
V3	64,218	64,420	0	45,722	56,104	0	57,020	59,762	0	37,908	49,188	0	6,551	11,768	2,110
V4	66,546	62,706	0	46,130	55,298	0	55,746	60,754	0	35,728	51,435	0	6,555	11,763	2,111
V5	67,224	60,119	0	46,094	53,193	0	57,207	61,401	0	35,624	53,866	0	6,205	11,729	2,110
V6	67,813	60,395	0	45,804	53,531	0	56,946	61,324	0	36,101	52,768	0	5,727	12,253	2,110
V7	67,426	60,394	0	45,730	52,945	0	56,369	62,204	0	36,410	52,576	0	60,11	12,624	2,083
Iteration 2															
V1	66,819	60,565	0	45,590	52,771	0	56,844	62,297	0	36,503	52,757	0	6,001	12,542	2,084
V2	63,656	65,152	0	48,291	53,574	0	56,490	60,242	0	37,346	49,592	0	6,572	11,747	2,110
V3	63,290	65,348	0	47,590	54,236	0	55,960	60,823	0	37,314	49,782	0	6,201	12,241	1,987
V4	63,540	65,711	0	47,641	53,787	0	55,875	60,625	0	37,309	49,854	0	6,238	12,099	2,092
V5	63,858	65,054	0	47,788	53,165	0	56,850	60,373	0	37,647	49,564	0	6,211	12,151	2,110
V6	63,990	64,912	0	47,757	53,166	0	56,881	60,395	0	37,640	49,635	0	6,123	12,163	2,110
V7	63,693	64,906	0	47,927	53,049	0	56,855	60,758	0	37,722	49,568	0	6,118	12,112	2,063
Iteration 3															
V1	63,634	64,883	0	47,903	53,036	0	56,910	60,767	0	37,748	49,609	0	6,119	12,100	2,064
V2	63,656	65,152	0	48,291	53,574	0	56,490	60,242	0	37,346	49,592	0	6,572	11,747	2,110
V3	63,541	65,097	0	48,101	53,725	0	56,294	60,489	0	37,401	49,696	0	6,446	11,926	2,058
V4	63,820	65,431	0	47,986	53,442	0	56,168	60,332	0	37,412	49,751	0	6,421	11,893	2,115
V5	63,908	65,239	0	48,045	53,241	0	56,467	60,258	0	37,537	49,647	0	6,397	11,923	2,110
V6	63,929	65,209	0	48,031	53,239	0	56,475	60,254	0	37,577	49,639	0	6,379	11,930	2,110
V7	63,810	65,306	0	48,046	53,193	0	564,53	60,382	0	37,617	49,590	0	6,372	11,927	2,077

The last row of Table A29 illustrates how close the raked marginal total weights are for the sample, compared to the estimated 2015 PUMS population total for Springfield.

Table A31 describes the maximum percent difference in marginal totals for the raked weights versus the population by raking variable and iteration. The largest percent difference in weights is for the raking variables age by collapsed raceV1 (V3), and age by education (V4). After 14 iterations, the maximum percent difference was minimized, with the maximum percent difference in cell weights between the PUMS totals and the Weighted sample less that 0.7%.

Table A31. Springfield: Max Percent Difference between Sample and Aligned Population Weights by Raking Iteration

Iteration	Max Difference	Race Max1	Age x Sex Max2	Age x RaceV1 Max3	Age x Edu Max4	Sex x RaceV1 Max5	Sex x Edu Max6	Rav2 x EduV1 Max7
1	38.15%	17.2%	8.5%	38.2%	16.1%	10.3%	2.4%	0.0%
2	14.47%	1.4%	6.9%	14.5%	9.9%	3.5%	3.5%	0.0%
3	8.060%	1.3%	3.1%	5.8%	8.1%	2.3%	2.3%	0.0%
4	6.280%	0.5%	1.2%	2.5%	6.3%	1.5%	1.5%	0.0%
5	4.862%	0.5%	0.8%	1.2%	4.9%	1.1%	1.1%	0.0%
6	3.767%	0.5%	0.6%	0.7%	3.8%	0.8%	0.8%	0.0%
7	2.933%	0.5%	0.4%	0.7%	2.9%	0.6%	0.6%	0.0%
8	2.300%	0.5%	0.4%	0.7%	2.3%	0.4%	0.4%	0.0%
9	1.819%	0.5%	0.4%	0.7%	1.8%	0.4%	0.4%	0.0%
10	1.454%	0.5%	0.4%	0.7%	1.5%	0.4%	0.3%	0.0%
11	1.177%	0.5%	0.4%	0.7%	1.2%	0.4%	0.2%	0.0%
12	0.9663%	0.5%	0.4%	0.7%	1.0%	0.4%	0.2%	0.0%
13	0.8058%	0.5%	0.4%	0.7%	0.8%	0.4%	0.2%	0.0%
14	0.6839%	0.5%	0.4%	0.7%	0.7%	0.4%	0.1%	0.0%
15	0.6839%	0.5%	0.4%	0.7%	0.6%	0.4%	0.1%	0.0%
16	0.6838%	0.5%	0.4%	0.7%	0.5%	0.4%	0.1%	0.0%
17	0.6838%	0.5%	0.4%	0.7%	0.5%	0.4%	0.1%	0.0%
18	0.6838%	0.5%	0.4%	0.7%	0.4%	0.4%	0.1%	0.0%
19	0.6838%	0.5%	0.4%	0.7%	0.4%	0.4%	0.1%	0.0%
20	0.6837%	0.5%	0.4%	0.7%	0.4%	0.4%	0.1%	0.0%
21	0.6837%	0.5%	0.4%	0.7%	0.3%	0.4%	0.1%	0.0%
22	0.6837%	0.5%	0.4%	0.7%	0.3%	0.4%	0.1%	0.0%
23	0.6837%	0.5%	0.4%	0.7%	0.3%	0.4%	0.1%	0.0%
24	0.6837%	0.5%	0.4%	0.7%	0.3%	0.4%	0.1%	0.0%
25	0.6837%	0.5%	0.4%	0.7%	0.3%	0.4%	0.1%	0.0%

As an illustration, the largest percentage difference in weights for V3 (age by collapsed race) can be examined after 14 iterations of raking in Table A32. The difference occurs for respondents aged 18-34 who were Hispanic or Asian, where the percent difference is given by $0.71\% = 100\left(\frac{8{,}904 - 8{,}859}{8{,}859}\right)$. The low percentage differences in marginal weights between the MA population estimates and the raked weight marginal totals indicate that by using the weights, the respondents closely mirror the MA population in the Springfield region by the raking variables.

Table A32. Springfield Baseline Targeted Survey: Sample and Aligned Population Weights After 14 Iterations of Raking

V1	Race				
	Hispanic	White	Black	Asian	Missing
PUMS	73,742	324,767	2,6868	10,287	19,109
Sample	74,077	324,418	2,6882	10,287	19,109

V2	Age/Gender														
	18-34			35-49			50-64			65+			Missing		
	Male	Female	Missing	Male	Female	Missing	Male	Female	Missing	Male	Female	Missing	Male	Female	Missing
PUMS	63,656	65,152	0	48,291	53,574	0	56,490	60,242	0	37,346	49,592	0	6,572	11,747	2,110
Sample	63,870	65,398	0	48,106	53,342	0	56,366	60,119	0	37,440	49,701	0	6,570	11,752	2,108

V3	Age/Race											
	18-34				35-49				50-64			
	White	Black	Hispanic/Asian	Missing	White	Black	Hispanic/Asian	Missing	White	Black	Hispanic/Asian	Missing
PUMS	35,483	79,902	8,859	4,394	24,180	68,783	7,366	1,496	14,495	92,133	6,507	3,647
Sample	35,726	80,226	8,904	4,412	24,143	68,475	7,341	1,490	14,499	91,854	6,494	3,638

V3 (cont.)	Age/Race							
	65+				Missing			
	White	Black	Hispanic/Asian	Missing	White	Black	Hispanic/Asian	Missing
PUMS	6,427	73,910	3,850	2,909	3,540	99,41	287	6,662
Sample	6,450	73,927	3,855	2,908	3,546	9,936	287	6,661

V4	Age/Race											
	18-34				35-49				50-64			
	<= High School	College	Graduate School	Missing	<= High School	College	Graduate School	Missing	<= High School	College	Graduate School	Missing
PUMS	53,074	68,580	7,272	326	39,341	48,775	12,987	326	48,926	52,984	13,451	1,140
Sample	53,092	68,569	7,282	326	39,372	48,746	13,002	328	48,935	52,945	13,466	1,139

V4 (cont.)	Age/Race							
	65+				Missing			
	<= High School	College	Graduate School	Missing	<= High School	College	Graduate School	Missing
PUMS	46,248	27,885	10,728	2,302	1,698	9,624	6,888	2,219
Sample	46,236	27,862	10,738	2,306	1,697	9,623	6,895	2,215

V5	Gender/Race											
	Male				Female				Missing			
	White	Black	Hispanic/Asian	Missing	White	Black	Hispanic/Asian	Missing	White	Black	Hispanic/Asian	Missing
PUMS	37,831	153,608	13,660	7,255	46,177	170,621	13,215	10,294	0	550	0	1,560
Sample	37,989	153,443	13,665	7,255	46,375	170,424	13,217	10,296	0	551	0	1,557

V6	Gender/Education											
	Male				Female				Missing			
	<= High School	College	Graduate School	Missing	<= High School	College	Graduate School	Missing	<= High School	College	Graduate School	Missing
PUMS	97,987	91,373	20,475	2,557	91,539	116,205	30,330	2,197	0	0	550	1,560
Sample	97,883	91,422	20,487	2,559	91,448	116,322	30,346	2,196	0	0	551	1,557

V7	Race/Education											
	Hispanic			White/Other			Black/Asian			Missing		
	<= High School	> High School	Missing	<= High School	> High School	Missing	<= High School	> High School	Missing	<= High School	> High School	Missing
PUMS	479,48	25,046	1,083	120,738	203,074	607	16,626	20,543	0	4,020	10,465	4,623
Sample	47,948	25,046	1,083	120,738	203,074	607	16,626	20,543	0	4,020	10,465	4,623

Step 6 Trimming Raked Weights

Weights developed via this process can have a broad range for individual respondents. While the weights provide in theory a way of obtaining unbiased estimates of population parameters, variability in weights will inflate the variance of the estimates. For this reason, trimming the weights can be desirable to improve the overall estimation accuracy. Trimming was examined in the BGPS weight development. We apply the strategy for trimming weights developed in the BGPS to weights developed for Springfield.

Trimming Raked Weights

We describe the procedure for trimming raked weights next. Let $w_{\min}$ represent the minimum weight, w_{mean} represent the mean weight, and $w_{\max}$ represent the maximum weight. We define the trimmed weight by setting the minimum and maximum weight to be a simple multiplier, m, times the average weight, w_{mean}. The initial trimmed weight is given by

$$w_{j,\mathrm{m}}^{0} = \begin{cases} w_{\max,m} & \text{if } w_j \ge w_{\max,m} \\ w_i & \\ w_{\min,m} & \text{if } w_j \le w_{\min,m} \end{cases} .$$

where $w_{\max,m} = m\left(w_{mean}\right)$ and $w_{\min,m} = \left(w_{mean}\right)/m$. By changing the minimum and maximum weight, the total weight is changed. In order to insure that the total weight is equal to the total population size, we adjust the initial trimmed weight by a factor $\frac{T}{T_m}$, where $T = \sum_{j=1}^{n} w_j$ represents the total raked weight prior to trimming, and $T_m = \sum_{j=1}^{n} w_{j,\mathrm{m}}^{0}$ represents the total weight after trimming weights to a multiple of the mean weight. The final step in forming the trimmed weight is to multiply the initial trimmed weight by $\frac{T}{T_m}$, to form the trimmed weight

$$w_{j,m} = \left(\frac{T}{T_m}\right) w_{\mathrm{j,m}}^{0} .$$

Determining the Extent of Trimming

In the BGPS, we determined the multiplier used to trim weights by evaluating the accuracy of estimators for values of $m = 2,3,4,5,6,7,8,9$ for the variables defined as a) problem gambler; b) at risk gambler; c) tobacco user; and d) participant in extreme sports. An unbiased estimator of the variable was assumed to be the estimator without trimming. Using this process, we found that the most accurate estimator will occur when $m = 8$. We use this multiplier to form trimmed weights.

Trimming Raked Weights for Springfield

By setting $m = 8$, the minimum and maximum raked weights are given by 50.26 and 3216.78, respectively. This results in trimming weights for 8 respondents (Table A33).

Table A33. List of Weights Trimmed for Springfield

Observation	Springfield Household Size Weight: SP1WT4	Springfield Raked Weight Not Trimmed: SP1WT5	Initial Trimmed Raked Weight: WT6a
1	219.80	29.21	50.26
2	166.92	32.97	50.26
3	362.82	41.87	50.26
4	333.84	46.32	50.26
5	362.82	49.44	50.26
6	1277.88	4099.86	3216.78
7	1277.88	4711.57	3216.78
8	1135.44	8289.96	3216.78

The total weight prior to trimming is 454,772, while the total weight after trimming (but prior to adjusting) is 447,372. We define the raked trimmed weight by multiplying the trimmed weights by the factor $\frac{T}{T_m} = \frac{454,772}{447,372}$. The resulting weight, SP1WT6, is the final weight for the Springfield Baseline Targeted Survey.

Appendix A2: Weighting Procedures for Springfield Follow-up Target Population Survey

The Springfield target survey selected addresses from a stratified sampling frame of address defined by Springfield and the surrounding communities. The study had two sample releases, with release 1 in October 2019 and release 2 in December 2019. The targeted area for the study was Springfield, Massachusetts and surrounding towns, including Agawam, Chicopee, East Longmeadow, Hampden, Holyoke, Longmeadow, Ludlow, Northampton, Wilbraham, and West Springfield. To achieve adequate respondents in the demographic cells of interest, we included race- targeted list sample in combination with the USPS frame for complete coverage. This involved oversampling of Hispanic and African American households included in the listings, as well as stratification of census tracts by low and high white population areas. The race-targeted list samples were purchased from Marketing Systems Group (MSG) and created from various publicly-available and modeled sources (surnames, purchases, etc.). The Census tracts were stratified by percent non-White population counts using data from the American Community Survey (ACS). The distribution of the percent non-White was examined to form the strata. The census tracts with the highest percent non-White formed one stratum, and the remaining tracts formed the other stratum. It high percent non-White stratum was defined was defined to include 20% of the tracts. The tracts selected were those that had the highest percent non-White population. We refer to this stratum as the High Non-White stratum. The remaining tracts are in the Low Non-White stratum.

The MSG race-targeted list frame contained 36,801 Hispanic likely addresses and 14,282 African American likely addresses. The remainder of the frame came from the residential addresses only on the USPS computerized delivery sequence (CDS) file, including PO box addresses if designated as 'only way to get mail' (OWGM) and drop delivery addresses with synthetic apartment numbers appended. After matching the list sample to the CDS, the frame for release 1 contained 170,486 total addresses (shown in Table A34).

Table A34. Frame Counts by Stratum

Stratum	Frequency	Percent
CDS-Only, Low Non-White	104,137	61%
CDS-Only, High Non-White	15,266	9%
List – Hispanic Only	36,801	22%
List – African American Only	14,282	8%

All list sample lines selected by MSG were included in the sample in addition to CDS addresses selected through systematic sampling across the target area with a geographic sort. Release 1 contained 4,689 sample lines and release 2 contained 3,510. Ultimately, we released 8,199 sampled addresses between the two releases. The CDS-only lines included 3,700 in the low stratum and 2,000 in the high or "oversample" stratum. The list sample included 1,722 targeted likely Hispanic addresses and 777 targeted likely African American addresses. Table A35 provides the breakdown of the sample by release and stratum.

Table A35. Sample by Release and Stratum

Stratum	Frequency Release 1	Frequency Release 2	Overall Count	Overall Percentage
CDS-Only, Low Non-White	1,774	1,926	3,700	45.1%
CDS-Only, High Non-White	968	1,032	2,000	24.4%
List – Hispanic Only	1,363	359	1,722	21.0%
List – Black Only	584	193	777	9.5%
All	4,689	3,510	8,199	100%

The stratified random sample of addresses included 8,199 addresses, with a sampling rate ranging from 3.55% in the CDS-Only, Low Non-White stratum to 13.10% in the CDS-Only, High Non-White stratum. The sampling rates were selected to achieve a minimum of 1000 completed surveys. Weights are assigned to survey respondents that make the weighted responses representative of the Springfield target community population. A list of census tracts assigned to the CDS strata is given in Table A36. A total of 65 census tracts were assigned to the CDS-Only, High-White stratum, while 28 census tracts were assigned to the CDS-Only, Low-White stratum.

Table A36. Census Tract Assigned to CDS Strata

CDS High-White Tract	CDS High-White Tract	CDS Low-White Tract
25013813500	25013810414	25013800400
25015821700	25013811302	25013802601
25015821904	25013813207	25013801503
25015821901	25013811200	25013811800
25015821903	25013812404	25013801700
25015822200	25013802602	25013800102
25015821602	25013810800	25013811700
25015821601	25013812300	25013801502
25015822000	25013812202	25013801402
25013810602	25013801604	25013800500
25013810412	25013810700	25013801501
25013813204	25013810902	25013801902
25013813208	25013812201	25013802200
25013813209	25013802400	25013802300
25013810404	25013802500	25013801901
25013813205	25013801603	25013801800
25013813601	25013810601	25013801200
25013810403	25013801601	25013801300
25013813401	25013810901	25013800800
25013813403	25013811101	25013801401
25013813301	25013811102	25013801101
25013813602	25013801602	25013800900
25013812403	25013812104	25013802000
25013811900	25013801605	25013811600

CDS High-White Tract	CDS High-White Tract	CDS Low-White Tract
25013813304	25013800202	25013811500
25013812101	25013812002	25013800600
25013812401	25013802100	25013811400
25013811000	25013800101	25013800700
25013813404	25013812001	
25013813206	25013800201	
25013811301	25013801102	
25013813303	25013812103	
	25013800300	

A total of 1,134 surveys were completed between 10/10/2019 and 2/17/2020. The simple response rates by stratum are given in Table A37, and ranged from 8.1% (for CDS-only addresses in the High Non-White stratum, *t=2*) to 17.95% (for CDS-only addresses in the Low Non-White stratum, *t=1*).

Table A37. Simple Response Rates by Stratum

Stratum	Response Complete		
	Sampled Addresses	Responses	Percent
CDS – Low Non-White	3,700	664	17.9%
CDS – High Non-White	2,000	162	8.1%
List – Hispanic Only	1,722	198	11.5%
List – Black Only	777	110	14.2%
All	8,199	1,134	13.8%

There are five steps in the weighing. First, a weight is assigned to directly account for the stratified sampling fraction of addresses. Next, using information on the sample addresses, adjustments were made for unknown eligibility of the addresses. We call this the 'screening' weight, since it accounts for differences in the percent of sampled addresses that can be screened. The third step in the weighting is to adjust for non-response. Household size is accounted for next. Finally, raking is used to align the weights with the distribution of four demographic variables (age, gender, race, and education) to the target population based on the 2018 American Community Survey PUMS data.

The weighting scheme involves the following steps:

1. Base sampling weight for sample of addresses;
2. Adjustment for unknown eligibility of addresses (screening);
3. Adjustment for non-response to the questionnaire;
4. Adjustment for household size;
5. Raking of weights to age, gender, race, and education based on the 2018 PUMS.

Each individual weighting step is discussed in detail below.

Step 1. Base sampling weight

The base weight reflects the probability of an address being selected from the population frame. A total of 8,199 addresses were selected via stratified systematic random sampling from the 170,486 addresses in the frame defined by NORC, resulting in a probability of selection that ranged from 3.55% to 13.10% (Table 38). The first weight, W_{t1} for stratum $t = 1,...,4$ is equal to the inverse of the probability of selection, and ranged from $W_{t1} = \frac{15,266}{2000} = 7.63$ for $t = 2$ (i.e. CDS-only, High Non-White) to $W_{t1} = \frac{104,137}{3,700} = 28.15$ for $t = 1$ (i.e. CDS-only, Low Non-White). The variable representing this weight is ST_WT1.

Table A38. Simple Sampling Rates by Stratum and Base Weight ST_WT1

Stratum (t)	Address Total ADDTOT	Sample Addresses SAMPTOT	Sampling Fraction SFRACT	Target Base Weight ST_WT1
	Mean	Mean	Mean	Mean
CDS – Low Non-White	104,137	3,700	0.0355	28.15
CDS – High Non-White	15,266	2,000	0.1310	7.63
List – Hispanic Only	36,801	1,722	0.0468	21.37
List – Black Only	14,282	777	0.0544	18.38

Step 2. Adjustment for Screened Address

The first adjustment to the weights consists of an adjustment to account for those cases that were unable to be contacted or sufficiently screened, and thus have an unknown eligibility status. In this adjustment, the weight for cases of known eligibility status is adjusted to account for cases with unknown eligibility status. This assumes that the unknown cases would be eligible in the same proportion as the cases of known status, taking into account various other factors.

Among the sample addresses, a total of 1,188 were contacted/sufficiently screened so that they could be assessed as to the eligibility of the address. Thus, the unknown eligibility rate,

$$\pi_2 = \frac{Known\_Eligiblity}{(Known\_Eligiblity) + (Unknown\_Eligibility)}$$

is estimated by $\hat{\pi}_2 = \frac{1,188}{8,199} = 0.1449$.

Some additional information was available from the address frame so that it was possible to examine differences in the eligibility rate between addresses with different characteristics. Characteristics examined included the basic stratification variables for the frame, the type of address (single family dwelling, multi-family dwelling), and the language (address in English speaking area, address in Spanish speaking area). An address was deemed “Spanish language area” if the area (specifically Census defined block group) had 25% or more of the households where Spanish was spoken. The screening rates by a cross-classification of the address characteristics are given in Table A39.

Table A39. Screening Rates by Type of Address, Dwelling Unit, and Language

Stratum	Dwelling Unit	Language	Screened		Weight ST-WT1	Weight ST-WT2
			N	Percent	Mean	Mean
CDS – Low Non-White	Single Family	English	2,658	20.1%	28.15	139.83
		Spanish	171	8.2%	28.15	343.77
	Multiple Family	English	797	16.1%	28.15	175.25
		Spanish	74	10.8%	28.15	260.34
	All		3,700	18.5%	28.15	152.02
CDS – High Non-White	Single Family	English	868	10.5%	7.63	72.81
		Spanish	136	8.1%	7.63	94.37
	Multiple Family	English	805	8.4%	7.63	90.36
		Spanish	191	3.7%	7.63	208.27
	All		2,000	8.9%	7.63	86.25
List – Hispanic Only	Single Family	Spanish	1,116	11.8%	21.37	180.68
	Multiple Family	Spanish	606	12.5%	21.37	170.41
	All		1,722	12.1%	21.37	176.93
List – Black Only	Single Family	English	622	15.9%	18.38	115.48
	Multiple Family	English	155	12.3%	18.38	149.95
	All		777	15.2%	18.38	121.03
All			8,199	14.5%	20.79	143.51

Table A39 summarizes the screening rates (for eligibility) by address groups. We use these rates to adjust weights for screening. We estimate the screening rate for each group, $k=1,\ldots,K=12$, as $\hat{\pi}_{2k}=\frac{m_{2k}}{n_{2k}}$, where m_{2k} is the number of screened sampled addresses (where eligibility was known) in group k and n_{2k} is the number of sampled addresses in group k. We use the estimated screening rate to define the weight, ST_WT2, adjusted for screening, given by

$$W_{2k}=\frac{W_1}{\hat{\pi}_{2k}} .$$

This weight is assigned to all 1,188 screened addresses where eligibility is known, and is set to 'missing' for addresses with unknown eligibility status.

Step 3 Adjustment for Interview Completion Rate

The target respondent from each screened address was attempted to be interviewed, and the interview was classified as complete or not complete. A total of 1,134 interviews were complete, yielding a completion rate of $\hat{\pi}_3=\frac{1{,}134}{1{,}188}=0.9545$.

It was possible to examine differences in the completion rate using variables for the address strata, type of address (single family dwelling, multi-family dwelling), and language (address in English speaking area, address in Spanish speaking area). We use the completion rates in a cross-classification of these variables to adjust the weights to account for the complete responses. The cross-classification is the same as was used for screening rates.

We estimate the completion rates for each group, $k=1,\ldots,K=12$, as $\hat{\pi}_{3k}=\frac{m_{3k}^{*}}{m_{2k}}$, where m_{3k}^{*} is the number of complete respondents among the screened respondents, m_{2k}, in group *k*. We use the estimated completion rate to define the weight adjusted for screening, and non-response given by

$$W_{3k} = \frac{W_{2k}}{\hat{\pi}_{3k}}$$

for $k = 1,...,1,134$ completed survey respondents. This weight is assigned to all 1,134 completed survey respondents, and is set to 'missing' for addresses where the survey was not complete. The results are summarized in Table A40.

Table A40. Completion Rates by Type of Address, Dwelling Unit, and Language

Stratum	Dwelling Unit	Language	Survey Completed		Target Base Weight ST_WT1	Target Screened Weight ST_WT2	Target Completion Weight ST_WT3
			N	Percent	Mean	Mean	Mean
CDS – Low Non-White	Single Family	English	535	96.6%	28.15	139.83	144.70
		Spanish	14	100.0%	28.15	343.77	343.77
	Multiple Family	English	128	97.7%	28.15	175.25	179.45
		Spanish	8	100.0%	28.15	260.34	260.34
	All		685	96.9%	28.15	152.02	156.83
CDS – High Non-White	Single Family	English	91	89.0%	7.63	72.81	81.80
		Spanish	11	100.0%	7.63	94.37	94.37
	Multiple Family	English	68	92.6%	7.63	90.36	97.53
		Spanish	7	100.0%	7.63	208.27	208.27
	All		177	91.5%	7.63	86.25	94.23
List – Hispanic Only	Single Family	Spanish	132	93.9%	21.37	180.68	192.34
	Multiple Family	Spanish	76	97.4%	21.37	170.41	175.01
	All		208	95.2%	21.37	176.93	185.86
List – Black Only	Single Family	English	99	93.9%	18.38	115.48	122.93
	Multiple Family	English	19	89.5%	18.38	149.95	167.59
	All		118	93.2%	18.38	121.03	129.84
All			1,188	95.5%	20.79	143.51	150.34

Step 4 Adjustment for Household Size

The fourth adjustment in the weights is for household size. The number of persons 18 years or older living the household was recorded for most (1,088 or 95.9%) of the completed interviews. The distribution of household size for completed respondents is given in Table A41, where we assigned a household size of "1" to respondents who reported a household size of zero (n=28), and re-assigned the household size to missing for two respondents who reported household sizes greater than 60 persons.

Table A41. Number of 18+ Persons in Respondent Households

# 18+ in Household	Frequency	Cumulative Percent	Cumulative Frequency	Percent
-	46	4.06%	46	4.06%
1	342	30.16%	388	34.22%
2	521	45.94%	909	80.16%
3	151	13.32%	1060	93.47%
4	51	4.50%	1111	97.97%
5	16	1.41%	1127	99.38%
6	3	0.26%	1130	99.65%
7	2	0.18%	1132	99.82%
9	1	0.09%	1133	99.91%
10	1	0.09%	1134	100.00%

Let the household size (based on 18+ age household members) for respondent be represented by p_j. The total number of addresses in the Springfield frame is estimated for the 1,134 respondents as $\sum_{j=1}^{1134} W_{3j} = 170,486$. This total can be divided into addresses where the number of 18+ persons in the household is known, $\sum_{j=1}^{1088} W_{3j} = 163,648$, and addresses where the number of 18+ persons is unknown, $\sum_{j^*=1}^{46} W_{3j^*} = 6,838$. Using the household size when it is known, the total number of persons in the 163,648 households is estimated by $\sum_{j=1}^{1088} W_{3j} p_j = 325,741$, resulting in an average household size of $\frac{325,741}{163,647} = 1.99$, which we round to 2. We assign this household size to the 46 respondents where household size is unknown. We further truncate the household size to a maximum of 4 in an effort to limit the variability of the survey weights. The weight adjusted for household size is given by

$$W_{4,j}^* = p_j W_{3,j}.$$

The total weight corresponds to the estimated number of 18+ persons in the target area, given by $\sum_{j=1}^{1134} W_{4j}^* = 333,257$. In the Springfield target area (defined by PUMA codes 1900, 1901, 1902, and 1600), the total number of persons age 18+ based on the 2018 PUMS data is 460,991. We adjust the total number of 18+ persons to match the PUMS total to define the household size adjusted weight given by

$$W_{4,j} = \frac{460,991}{333,257} W_{4,j}^*.$$

The weights are summarized in Table A42.

Table A42. Respondent Weights by Type of Address, Dwelling Unit, and Language, and Number of Household Members

Addresses	Dwelling Unit	Language	Survey Completed	Target Base Weight ST_WT1	Target Screened Weight ST_WT2	Target Completion Weight ST_WT3	Target Household Size Weight ST_WT4
			N	Mean	Mean	Mean	Mean
CDS – Low Non-White	Single Family	English	517	28.15	139.83	144.70	407.68
		Spanish	14	28.15	343.77	343.77	1,086.94
	Multiple Family	English	125	28.15	175.25	179.45	369.38
		Spanish	8	28.15	260.34	260.34	585.21
	All		664	28.15	152.25	156.83	416.93
CDS – High Non-White	Single Family	English	81	7.63	72.81	81.80	240.26
		Spanish	11	7.63	94.37	94.37	261.09
	Multiple Family	English	63	7.63	90.36	97.53	241.99
		Spanish	7	7.63	208.27	208.27	576.20
	All		162	7.63	86.95	94.23	256.87
List – Hispanic Only	Single Family	Spanish	124	21.37	180.68	192.34	596.49
	Multiple Family	Spanish	74	21.37	170.41	175.01	405.67
	All		198	21.37	176.84	185.86	525.17
List – Black Only	Single Family	English	93	18.38	115.48	122.93	354.74
	Multiple Family	English	17	18.38	149.95	167.59	327.29
	All		110	18.38	120.81	129.84	350.50
All			1,134	23.08	144.17	150.34	406.52

Step 5 Aligning to 2018 Population Via Raking

We determine a set of raking variables after review of data from the 2018 ACS PUMS data set, and review of the survey respondent data. For the PUMS data, the review considered the adequacy of the PUMS data for raking cells by examining the number of respondents in a cell, and the coefficient of variation of the mean (CVM) weight (statistical weight assigned in the ACS) to a cell. When the number of respondents is small (less than 20), and/or the CVM is large (greater than 10%), some cells were combined so as to produce stable estimates of the total persons in the cell. When the CVM is less than 10%, we consider the uncertainty in estimating the total number of adults in the cell to be acceptably small. In this situation, we consider the assumption of ignoring the uncertainty in the population total weight estimates for each cell to be justified.

In addition to the investigation of the adequacy of data in cells from the PUMS, we examined the number of respondents in each cell in the targeted survey. Cells where there were fewer than 10 respondents were identified, and considered for possible collapsing.

Primary Variables for Raking Weights for Targeted Areas using the 2018 PUMS Population Data

We consider four variables: age (18-34, 35-49, 50-64, 65+), gender (male, female), race/ethnicity (Hispanic, Black (only), Asian (only), White/Other), and education (high school or less, some college/college graduate, some postgraduate education). We consider these variables individually and pair-wise. For some variables, we also consider a collapsed set of categories (i.e., for race/ethnicity: Hispanic/ Asian and Black/Asian, and for education: high school or less, or more than high school). We determined raking variables via a preliminary analysis of the 2018 one-year American Community Survey Public Use Microdata Sample (PUMS) files. In an ideal setting, reliable PUMS data for population totals would be available for a full cross-classification of weighting variables. In practice, estimates of the population based on the PUMS data are based on an approximate 1% sample of the MA population and the PUMS data themselves are weighted to estimate the number of individuals in each post-stratum. We evaluate the reliability of the PUMS estimates of totals in cells formed by individual variables and a cross-classification of pairs of the variables. Reliability was evaluated by: (1) the number of PUMS respondents for each possible raking cell and (2) calculating the CVM in the estimated total PUMS weight. The CVM is the estimated standard error of the mean, divided by the mean, and expressed as a percent.

Table A43 provides the CVM and the number of respondents in the 2018 PUMS for the Springfield target area for each of the four variables used in weighting. The CVM was less than 10% for all cells, and in all cells, there are greater than 20 respondents.

Table A43. CVM and N by Age, Gender, Race, and Education for Springfield Using 2018 PUMS Data

PUMS 2018	Age				Gender		Race			
	18-34	35-49	50-64	65+	Male	Female	Hispanic	White/Other	Black	Asian
CVM (Wt)	2.2%	2.1%	2.1%	1.8%	1.6%	1.4%	2.7%	1,1%	4.4%	8.9%
n	1,303	1,044	1,270	1,273	2,278	2,612	621	3,884	275	110

PUMS 2018	Education		
	<= High School	College	Graduate School
CVM (Wt)	1.8%	1.5%	3.0%
n	1,896	2,395	599

Next, we consider pair-wise cross-classifications of the weighting variables by region (Tables A44-A46). Once again, we highlight cells (in red) where the number of respondents is less than 20 and the CVM is greater than 10%.

Table A44. CVM and N by Two Variables (Age and one other) for Springfield Using 2018 PUMS Data

PUMS 2015			Age			
			18-34	35-49	50-64	65+
Gender	Male	CVM (Wt)	3.2%	3.3%	3.1%	2.7%
		n	601	512	607	558
	Female	CVM (Wt)	3.0%	2.7%	2.8%	2.5%
		n	702	532	663	715
Education	<= High School	CVM (Wt)	3.7%	3.9%	3.4%	2.6%
		n	428	356	506	606
	College	CVM (Wt)	2.8%	2.7%	2.9%	3.0%
		n	806	499	619	471
	Graduate School	CVM (Wt)	8.5%	4.6%	6.5%	5.0%
		n	69	189	145	196
Race	Hispanic	CVM (Wt)	4.2%	5.0%	6.5%	7.2%
		n	250	179	124	68
	White/Other	CVM (Wt)	2.6%	2.4%	2.1%	1.9%
		n	970	775	1,073	1,129
	Black	CVM (Wt)	8.2%	7.1%	9.7%	7.7%
		n	95	58	58	64
	Asian	CVM (Wt)	15.0%	9.6%	20.4%	19.9%
		n	51	32	15	12

Table A45. CVM and N by Two Variables (Race and one other) for Springfield Using 2018 PUMS Data

PUMS 2018			Race			
			Hispanic	White/Other	Black	Asian
Gender	Male	CVM (Wt)	4.1%	1.7%	6.0%	14.1%
		n	257	1,842	138	41
	Female	CVM (Wt)	3.5%	1.5%	6.5%	11.0%
		n	364	2,042	137	69
Education	<= High School	CVM (Wt)	3.3%	2.0%	6.4%	10.6%
		n	385	1,372	109	30
	College	CVM (Wt)	4.7%	1.5%	5.7%	15.0%
		n	213	1,977	141	64
	Graduate School	CVM (Wt)	17.3%	2.8%	21.9%	20.4%
		n	23	535	25	16

Table A46. CVM and N by Two Variables (Gender x Education) for Springfield Using 2018 PUMS Data

PUMS 2018			Education		
			<= High School	College	Graduate School
Gender	Male	CVM (Wt)	2.5%	2.5%	4.9%
		n	987	1,038	253
	Female	CVM (Wt)	2.4%	1.9%	3.8%
		n	909	1,357	346

For Springfield, the results in Tables A44-A46 indicate that using four categories for race will result in a large CVM for Asians when cross-classified by age (Table A44), gender (Table A45), and education (Table

A46). We examined the distribution of the variables age, gender, and education with race, in an effort to identify a racial group most similar to Asians. The results are given (using the weighted PUMS data) in Table A47.

Table A47. Estimated Totals by Race for Age, Gender, and Education (2018 PUMS Springfield)

PUMS 2018			Race				
			Hispanic	White/Other	Black	Asian	Total
Age	18-34	Frequency	38,034	84,592	94,47	4,516	136,589
		Column %	43.82%	25.44%	32.53%	35.69%	
	35-49	Frequency	24,212	69,199	7,424	3,251	104,086
		Column %	27.89%	20.81%	25.56%	25.69%	
	50-64	Frequency	16,334	94,121	7,456	3,263	121,174
		Column %	18.82%	28.31%	25.67%	25.79%	
	65+	Frequency	8,223	84,578	4,717	1,624	99,142
		Column %	9.47%	25.44%	16.24%	12.83%	
Total			86,803	332,490	29,044	12,654	460,991
Gender	Male	Frequency	40,136	156,357	14,027	5,808	216,328
		Column %	46.24%	47.03%	48.30%	45.90%	
	Female	Frequency	46,667	176,133	15,017	6,846	244,663
		Column %	53.76%	52.97%	51.70%	54.10%	
Total			86,803	332,490	29,044	12,654	460,991
Education	<= High School	Frequency	57,289	119,018	11,810	4,733	192,850
		Column %	66.00%	35.80%	40.66%	37.40%	
	College	Frequency	27,002	169,646	14,236	5,671	216,555
		Column %	31.11%	51.02%	49.02%	44.82%	
	Graduate School	Frequency	2,512	43,826	2,998	2,250	51,586
		Column %	2.89%	13.18%	10.32%	17.78%	
Total			86,803	332,490	29,044	12,654	460,991

The results in Table A47 suggest similar distributions of age for Blacks and Asians, similar distributions of gender with Hispanics and Asians, and more similar distributions of education with whites, particularly after collapsing the categories of college and graduate education. We created additional variables to reflect these combinations of race categories. One variable, P_RACEV1, combines Hispanics and Asians into a single group. A second variable, P_RACEV2, combines Blacks and Asians into a single group. A third variable, P_RACEV3, combines Whites and Asians into a single group. A fourth variable combines college and graduate education into a single group (P_EDUV1). Tables A48-A50 re-evaluates the CVM and number of subjects using these variables when cross-classified.

Table A48. CVM and N by Two Variables (Age and Collapsed Race) for Springfield Using 2018 PUMS Data

PUMS 2018			Age			
			18-34	35-49	50-64	65+
Race	Hispanic/Asian	CVM (Wt)	4.2%	5.0%	6.5%	7.2%
		n	250	179	124	68
	White/Other	CVM (Wt)	2.6%	2.4%	2.1%	1.9%
		n	907	775	1,073	1,129
	Black	CVM (Wt)	7.3%	5.8%	9.5%	8.2%
		n	146	90	73	76

Table A49. CVM and N by Two Variables (Gender and Collapsed Race) for Springfield Using 2018 PUMS Data

PUMS 2018			Gender	
			Male	Female
Race	Hispanic/Asian	CVM (Wt)	4.0%	3.4%
		n	298	433
	White/Other	CVM (Wt)	1.7%	1.5%
		n	1,842	2,042
	Black	CVM (Wt)	6.0%	6.5%
		n	138	137

Table A50. CVM and N by Two Variables (Collapsed Education and Collapsed Race) for Springfield Using 2018 PUMS Data

PUMS 2018			Education	
			<= High School	> High School
Race	Hispanic/Asian	CVM (Wt)	3.3%	4.5%
		n	385	236
	White/Other	CVM (Wt)	2.0%	1.4%
		n	1,402	2,592
	Black	CVM (Wt)	6.4%	6.0%
		n	109	166

For Springfield, after collapsing race, and education as indicated in Tables A48-A50 above, all cells contain at least 20 respondents and have a CVM less than 10%.

Respondents in the Springfield Targeted Survey

In addition to having stable estimates of the population total in each raking cell, we examine the number of respondents in the targeted survey by possible raking variables. If there are few respondents in a particular category, the weight assigned to the subjects may be difficult to match to the population weights. When there are no respondents in a cell, the cell needs to be collapsed to match population weights. Table A51 considers each raking variable individually. Note that there are more than 10 respondents in each non-missing cell.

Table A51. Number of Respondents by Demographics for Springfield Targeted Survey 2019

Respondents 2019 Survey		n
Age	18-34	246
	35-49	254
	50-64	269
	65+	284
	Missing	81
Gender	Male	406
	Female	710
	Missing	18
Education	<= High School	224
	College	615
	Graduate School	254
	Missing	41
Race	Hispanic	270
	White/Other	708
	Black	70
	Asian	21
	Missing	65

The results in Table A51 indicate that there are greater than 10 respondents in each demographic cell in the targeted survey (except for cells where the variable was missing). We next considered two way cross-classifications of variables, and examined cells to see if there were at least 10 respondents. We used the variables with collapsed categories that were obtained from the 2018 PUMS data. The resulting cross-classifications are given in Tables A52-A54.

Table A52. Number of Respondents by Demographics for Springfield Targeted Survey 2019

Respondents 2019 Survey		Age				
		18-34	35-49	50-64	65+	Missing
		n	n	n	n	n
Gender	Male	76	79	96	129	26
	Female	167	175	173	155	40
	Missing	3	-	-	-	15
Race	Hispanic/Asian	94	81	51	23	21
	White/Other	113	143	189	236	27
	Black	28	23	19	17	4
	Missing	11	7	10	8	29
Education	<= High School	53	39	59	57	16
	College	149	141	149	146	30
	Graduate School	37	71	58	78	10
	Missing	7	3	3	3	25

Table A53. Number of Respondents by Demographics for Springfield Targeted Survey 2019

Respondents 2019 Survey		Gender		
		Male	Female	Missing
		n	n	n
Race	Hispanic/Asian	90	200	1
	White/Other	280	420	8
	Black	20	50	-
	Missing	16	40	9
Education	<= High School	85	139	-
	College	220	392	3
	Graduate School	94	159	1
	Missing	7	20	14

Table A54. Number of Respondents by Demographics for Springfield Targeted Survey 2019

Respondents 2019 Survey		Education		
		<= High School	> High School	Missing
		n	n	n
Race	Hispanic/Asian	98	165	7
	White/Other	96	622	11
	Black	20	50	-
	Missing	10	32	23

The results in Tables A52-A54 indicate that there are greater than 10 respondents in each demographic cell in the two-way classifications when variables are collapsed similar to those in the PUMS data.

Raking Weights for the Springfield Targeted Survey using the 2015 ACS Population Data

We adjusted weights assigned to subjects to more closely align with the distribution of 18+ year old persons in MA for Springfield using a raking procedure. Based on data from the 2018 American Community Survey Public Use Microdata Sample (PUMS) files, and the distribution of respondents to the

Springfield targeted survey, adjustments were made for age (18-34, 35-49, 50-64, 65+), gender (male, female), race/ethnicity (Hispanic, Black (only), Asian (only), White and Other), education (high school or less, some college/college graduate, some post graduate education), and cross-classifications of variables as indicated in Table A50. For cross-classifications that included race and age or gender, categories for Hispanics and Asians are combined. For cross-classifications that included race and education, categories for Blacks and Asians are combined, and categories for college and graduate education are combined. The weights are developed are based on raking using a similar procedure as was used for developing raked weights in the Springfield Baseline Targeted Survey. The raking variables used are V1-V7 as indicated in Table A55.

Table A55. Variables using in Raking for Springfield Targeted Survey

Raking Variable	V1	V2	V3	V4	V5	V6	V7
Variables (# of Cells)	Race (5)	Age x Gender (15)	Age x Race (20)	Age x Edu (20)	Gender x Race (12)	Gender x Edu (12)	Race x Edu (12)
1st Variable Name	race_ps	age_ps	age_ps	age_ps	sex_ps	sex_ps	racec3_ps
Categories 1st Variable	Hispanic	18-34	18-34	18-34	Male	Male	Hispanic
	White/Other	35-49	35-49	35-49	Female	Female	White/Other
	Black	50-64	50-64	50-64	Missing	Missing	Black
	Asian	65+	65+	65+			Asian
	Missing	Missing	Missing	Missing			Missing
Categories 1st Variable		sex_ps	racec2_ps	edu_ps	racec1_ps	edu_ps	educ1_ps
		Male	Hispanic	High School or less	Hispanic	Hispanic	High School or less
		Female	White/Other	College	White/Other	White/Other	Some College
		Missing	Black/Asian	Graduate School	Black/Asian	Black/Asian	Missing
			Missing	Missing	Missing	Missing	

We first account for missing values for race, age x gender, age x collapsed race, age x education, gender x collapsed race, gender x education, and collapsed race x collapsed education by assigning weights to respondents with some missing values equal to the weights in the Springfield targeted survey. This process is explained in more detail below.

Accounting for Missing Data among Respondents for Raking Variables

Several steps are taken to develop raked weights. First, a weight is assigned to respondents who are missing response for each of the raked variables. Next, the total weight assigned to categories of a variable for respondents is matched to the total weight in the 2018 PUMS data.

We begin assuming that weights have been developed for respondents in the Springfield targeted survey such that the total weight for respondents matches the total adult population defined by the PUMA codes for the Springfield area in 2018 (i.e. 460,991).

This weight is SP_WT4. Seven variables are considered for raking, with some variables defined by a cross-classification of two other variables. For each variable, we add a missing value category. Then, using the augmented categories, we create a variable that uniquely defines the age (5) x gender (3) x Race (5) x Education (4) x Collapsed Education (3) x Collapsed RaceC1 (4) x Collapsed RaceC2 (4) x

Collapsed RaceC3 (4) levels. This variable has 57,600 categories. A similar variable is created using the 2018 PUMS data for the area. Using these categories, the total weight (SP_WT4) is summed for each category using the Springfield data, and using the PUMS data (using the pwgpt variable). These two sets of counts (each totaling the 2018 population total) are the input for the raking.

Raking is an iterative process, with 1 iteration corresponding to an attempt to match marginal for each of the tables. For Springfield, we begin with the 7 raking variables listed in Table A56.

Table A56. List of Raking Variables and Number of Categories (including Missing Categories) for Raking in Springfield

	Variables	Categories
V1	Race/Ethnicity	5
V2	Age x Gender	15
V3	Age x Collapsed RaceC2	15
V4	Age x Education	20
V5	Gender x Collapsed RaceC1	12
V6	Gender x Education	12
V7	Collapsed RaceC3 x Collapsed EducationC1	12

We use the indices $i=1,...,5$ for categories of age (including 'missing age' as $i=5$), $j=1,...,3$ for categories of gender (including 'missing gender' as $j=3$), $k=1,...,5$ for categories of race/ethnicity (including 'missing race/ethnicity' as $k=5$), $l=1,...,4$ for categories of education (including 'missing education' as $l=4$), $m=1,...,3$ for categories of collapsed education (including 'missing collapsed education' as $m=3$), $n=1,...,4$ for categories of collapsed racec1 (including 'missing collapsed race' as $n=4$), $p=1,...,4$ for categories of collapsed racec2 (including 'missing collapsed race' as $p=4$) and $q=1,...,4$ for categories of collapsed racec3 (including 'missing collapsed race' as $q=4$).

We illustrate the process of raking weights using the first raking variable, V1, corresponding to race/ethnicity, and then describe the overall raking process. Raking was accomplished using a SAS program written for this purpose. The first step was to evaluate the total weight (NSP1WT4) in each of the 5 cells for the sample. Let us refer to these weights by x_k for $k=1,...,5$ (corresponding to race, including '5' as a missing age category), so that the total sample weight is given by $x_+ = \sum_{k=1}^{5} x_k$, where $x_k = \sum_{i=1}^{5}\sum_{j=1}^{3}\sum_{l=1}^{4}\sum_{m=1}^{3}\sum_{n=1}^{4}\sum_{p=1}^{4}\sum_{q=1}^{4} x_{ijklmnpq}$. The population weights, p_k, are based on the 2018 PUMS data. Among the population data, there are no missing values. Using the categories of race, the total population is the sum over 4 race cells, $p_+ = \sum_{k=1}^{4} p_k$, where $p_k = \sum_{i=1}^{5}\sum_{j=1}^{3}\sum_{l=1}^{4}\sum_{m=1}^{3}\sum_{n=1}^{4}\sum_{p=1}^{4}\sum_{q=1}^{4} p_{ijklmnpq}$. As a result, when raking by the variable V1, we first re-allocated PUMS data to form categories representing "missing age."

Table A57 illustrates these totals for Springfield prior to and after adjusting for missing race/ethnicity data.

Table A57. Initial Weights and Missing-Adjusted PUMS Weights for Race for Springfield (2018 PUMS) V1

Categories		Springfield Wt	PUMS Wt	Rev PUMS Wt
		Sum	Sum	Sum
Race	Hispanic	119,409	86,803	81,720
	White/Other	279,517	332,490	313,020
	Black	24,415	29,044	27,343
	Asian	10,657	12,654	11,913
	Missing	26,995	0	26,995
	All	460,992	460,991	460,991

The population weight for the missing data is calculated by assigning population weights to cells where race is missing proportional to the weight assigned these cells in the sample, $p_5^* = p_+\left(\frac{x_5}{x_+}\right)$. Since $p_+ = x_+$ (except for differences due to rounding), $p_5^* = x_5 = 26{,}995$ (representing weights assigned to the population with missing race). Weights for other age categories are adjusted, to preserve the overall population weight, p_+, such that $p_i^* = p_i\left(\frac{p_+ - p_5^*}{p_+}\right)$, for $i = 1,\ldots,4$. For example, the adjusted 2015 PUMS weight for Hispanics in Springfield is given by $81{,}720 = 86{,}803\left(\frac{460{,}991 - 26{,}995}{460{,}991}\right)$. Similar results are given for each of the other raking variables (Tables A58-A63).

Table A58. Weights Accounting for Missing Values for Age x Gender for Springfield Follow-up (2018 PUMS) V2

Age Categories		Age					
		18-34	35-49	50-64	65+	Missing	All
		Total Wt Sum	Total Wt Sum	Total Wt Sum	Total Wt Sum	Total Wt Sum	Total Wt Sum
PUMS 2018	Male	62,283	46,188	54,569	39,700	9,874	212,614
	Female	63,259	50,134	57,588	51,926	17,597	240,503
	Missing	976	0	0	0	6,897	7,874
	All	126,518	96,321	112,157	91,626	34,368	460,991
Springfield	Male	36,208	33,639	39,990	45,164	9,874	164,875
	Female	70,089	76,852	72,083	51,623	17,597	288,244
	Missing	976	0	0	0	6,897	7,874
	All	107,274	110,491	112,073	96,787	34,368	460,992

Table A59. Weights Accounting for Missing Values in V3: Age x RaceC2 for Springfield Follow-up (2018 PUMS)

Race Categories		Age					
		18-34	35-49	50-64	65+	Missing	All
		Total Wt Sum	Total Wt Sum	Total Wt Sum	Total Wt Sum	Total Wt Sum	Total Wt Sum
PUMS 2018	Hispanic	32,163	20,601	13,659	7,051	8,353	81,827
	White/Other	76,317	62,815	83,966	77,375	12,468	312,941
	Black/Asian	12,483	9,602	9,476	5,748	1,918	39,227
	Missing	4,705	3,026	5,423	2,212	11,629	26,995
	All	125,669	96,044	112,523	92,386	34,368	460,991
Springfield	Hispanic	43,365	35,859	21,631	10,201	8,353	119,409
	White/Other	48,191	61,938	77,901	79,020	12,468	279,517
	Black/Asian	11,013	9,669	7,119	5,354	1,918	35,071
	Missing	4,705	3,026	5,423	2,212	11,629	26,995
	All	107,274	110,491	112,073	96,787	34,368	460,992

Table A60. Weights Accounting for Missing Values in V4: Age x Edu for Springfield Follow-up (2018 PUMS)

Education Categories		Age					
		18-34	35-49	50-64	65+	Missing	All
		Total Wt Sum	Total Wt Sum	Total Wt Sum	Total Wt Sum	Total Wt Sum	Total Wt Sum
PUMS 2018	<= High School	48,219	38,657	46,514	45,617	6,155	185,163
	College	67,612	42,613	51,754	32,971	12,789	207,739
	Graduate School	7,073	13,719	12,120	12,344	4,405	49,661
	Missing	3,566	1,177	1,727	941	11,019	18,429
	All	126,469	96,165	112,115	91,873	34,368	460,991
Springfield	<= High School	23,440	15,455	26,066	19,138	6,155	90,253
	College	66,422	62,671	56,783	47,917	12,789	246,583
	Graduate School	13,847	31,188	27,496	28,791	4,405	105,727
	Missing	3,566	1,177	1,727	941	11,019	18,429
	All	107,274	110,491	112,073	96,787	34,368	460,992

Table A61. Weights Accounting for Missing Values in V5: Gender x RaceC1 for Springfield Follow-up (2018 PUMS)

Race Categories		Gender			
		Male	Female	Missing	All
		Total Wt Sum	Total Wt Sum	Total Wt Sum	Total Wt Sum
PUMS 2018	Hispanic/Asian	43,818	49,538	232	93,588
	White/Other	148,043	161,868	3,153	313,065
	Black	13,410	13,934	0	27,344
	Missing	7,360	15,146	4,488	26,995
	All	212,632	240,486	7,874	460,991
Springfield	Hispanic/Asian	40,180	89,654	232	130,065
	White/Other	110,013	166,351	3,153	279,517
	Black	7,322	17,093	0	24,415
	Missing	7,360	15,146	4,488	26,995
	All	164,875	288,244	7,874	460,992

Table A62. Weights Accounting for Missing Values in V6: Gender x Edu for Springfield Follow-up (2018 PUMS)

Education Categories		Gender			
		Male	Female	Missing	All
		Total Wt Sum	Total Wt Sum	Total Wt Sum	Total Wt Sum
PUMS 2018	<= High School	98,300	87,034	0	185,335
	College	90,290	116,281	1,180	207,752
	Graduate School	20,660	28,416	400	49,476
	Missing	3,448	8,687	6,293	18,429
	All	212,698	240,419	7,874	460,991
Springfield	<= High School	35,433	54,820	0	90,253
	College	85,310	160,093	1,180	246,583
	Graduate School	40,684	64,643	400	105,727
	Missing	3,448	8,687	6,293	18,429
	All	164,875	288,244	7,874	460,992

Table A63. Weights Accounting for Missing Values in V7: Racec3 x EduC1 for Springfield Follow-up (2018 PUMS)

Education Categories		Race				
		Hispanic	White/Other	Black/Asian	Missing	All
		Total Wt Sum	Total Wt Sum	Total Wt Sum	Total Wt Sum	Total Wt Sum
PUMS 2018	<= High School	53,059	116,816	11,338	3,486	184,699
	> High School	26,427	202,041	15,995	13,401	257,864
	Missing	2,903	5,417	0	10,108	18,429
	All	82,389	324,275	27,332	26,995	460,991
Springfield	<= High School	41,793	36,938	8,035	3,486	90,253
	> High School	74,712	247,818	16,379	13,401	352,310
	Missing	2,903	5,417	0	10,108	18,429
	All	119,409	290,174	24,415	26,995	460,992

The next step is to iteratively alter the respondent's weights so that the sample total weight for each raking table matches the population weight for the table. We illustrate this with an example. The totals for the Springfield Sample and the PUMS population differ for different categories of the raking variables. As an example from Table A63, the estimated number of Hispanics who had a High School or less education based on the Springfield sample is 41,793, while the estimated number from the Springfield PUMS is 53,059. The goal of raking is to alter the weights so that the Springfield sample cell totals are as similar as possible for the Springfield PUMS totals.

Description of a Step in the Raking

Raking is accomplished using a SAS program written for this purpose. The process proceeds in an iterative manner. Each iteration consists of a sequence of steps, where each step aligns the sample and population weights for a raking variable. We describe the process for the first raking variable, V1, corresponding to race/ethnicity.

The total sample weight assigned to a cell for a raking variable is the sum of ST_WT4 assigned to respondents in that cell. We index categories for the 4 primary variables by $i=1,\ldots,5$ for age, $j=1,\ldots,3$ for gender, $k=1,\ldots,5$ for race, $l=1,\ldots,4$ for education, $m=1,\ldots,3$ for collapsed education, $n=1,\ldots,4$ for collapsed racec1, $p=1,\ldots,4$ for collapsed racec2, and $q=1,\ldots,4$ for collapsed racec3. Respondents within a cell are indexed by $r=1,\ldots,n_{ijklmnpq}$. The total sample weight assigned to a cell for the first raking variable, V1 (i.e. race), is given by

$$x_k = \sum_{i=1}^{5}\sum_{j=1}^{3}\sum_{l=1}^{4}\sum_{m=1}^{3}\sum_{n=1}^{4}\sum_{p=1}^{4}\sum_{q=1}^{4} x_{ijklmnpq}$$
$$= \sum_{i=1}^{5}\sum_{j=1}^{3}\sum_{l=1}^{4}\sum_{m=1}^{3}\sum_{n=1}^{4}\sum_{p=1}^{4}\sum_{q=1}^{4}\left(\sum_{r=1}^{n_{ijklmnpq}} x_{ijklmnpqr}\right),$$

where $\sum_{r=1}^{n_{ijklmnpq}} x_{ijklmnpqr}$, and $x_{ijklmnpqr}$ is the value of ST_WT4 for respondent r . The first step in an iteration of raking aligns the sample marginal to the population marginal by forming the new weight for cells based on the full cross-classification of the raking variables, such that

$$x^{(1)}_{ijklmnpq} = x_{ijklmnpq}\left(\frac{p^*_k}{x_k}\right),$$

and p^*_k is the PUMS weight adjusted for missing data.

Using these weights, the total weight is evaluated for each cell for the raking variable, and for each cell corresponding to the remaining raking variables, such as V2 (corresponding to age x gender), etc, i.e. $x^{(1)}_{ij} = \sum_{k=1}^{5}\sum_{l=1}^{4}\sum_{m=1}^{3}\sum_{n=1}^{4}\sum_{p=1}^{4}\sum_{q=1}^{4} x^{(1)}_{ijklmnpq}$. Once again, using the population marginal weights, we align the sample marginal to the population marginal for V2, such that

$$x^{(2)}_{ijklmnpq} = x^{(1)}_{ijklmnpq}\left(\frac{p^*_{ij}}{x^{(1)}_{ij}}\right).$$

This process is continued for each of the raking variables, resulting in the marginal total weights in each cell after one iteration given by $r^1_{ijklmnpq} = x^{(7)}_{ijklmnpq}$.

Table A64 summarizes the sample and aligned population weights prior to a raking iteration, but after accounting for missing data. We note at the bottom of Table A64 that the maximum percent difference in population weights for any of the raking cells is 182.7%, which occurs for Hispanics with more than a high school education, i.e.

$$182.7\% = 100\left(\frac{74,712 - 26,427}{26,427}\right).$$

Table A64 illustrates the marginal weight totals for one variable, age x gender, on the 7 steps for the first three iterations of raking.

Table A64. Springfield 2019 Targeted Survey: Sample and Aligned Population Weights Prior to Raking

V1	Race				
	Hispanic	White	Black	Asian	Missing
PUMS	81,720	313,020	27,343	11,913	26,995
Sample	119,409	279,517	24,415	10,657	26,995

V2	Age/Gender														
	18-34			35-49			50-64			65+			Missing		
	Male	Female	Missing	Male	Female	Missing	Male	Female	Missing	Male	Female	Missing	Male	Female	Missing
PUMS	62,283	63,259	976	46,188	50,134	0	54,569	57,588	0	39,700	51,926	0	9,874	17,597	6,897
Sample	36,208	70,089	976	33,639	76,852	0	39,990	72,083	0	45,164	51,623	0	9,874	17,597	6,897

V3	Age/Race											
	18-34				35-49				50-64			
	White	Black/Asian	Hispanic	Missing	White	Black/Asian	Hispanic	Missing	White	Black/Asian	Hispanic	Missing
PUMS	32,163	76,317	12,483	4,705	20,601	62,815	9,602	3,026	13,659	83,966	9,476	5,423
Sample	43,365	48,191	11,013	4,705	35,859	61,938	9,669	3,026	21,631	77,901	7,119	5,423

V3 (cont.)	Age/Race							
	65+				Missing			
	White	Black/Asian	Hispanic	Missing	White	Black/Asian	Hispanic	Missing
PUMS	7,051	77,375	5,748	2,212	8,353	12,468	1,918	11,629
Sample	10,201	79,020	5,354	2,212	8,353	12,468	1,918	11,629

V4	Age/Race											
	18-34				35-49				50-64			
	<= High School	College	Graduate School	Missing	<= High School	College	Graduate School	Missing	<= High School	College	Graduate School	Missing
PUMS	48,219	67,612	7,073	3,566	38,657	42,613	13,719	1,177	46,514	51,754	12,120	1,727
Sample	23,440	66,422	13,847	3,566	15,455	62,671	31,188	1,177	26,066	56,783	27,496	1,727

V4 (cont.)	Age/Race							
	65+				Missing			
	<= High School	College	Graduate School	Missing	<= High School	College	Graduate School	Missing
PUMS	45,617	32,971	12,344	941	6,155	12,789	4,405	11,019
Sample	19,138	47,917	28,791	941	6,155	12,789	4,405	11,019

V5	Gender/Race											
	Male				Female				Missing			
	White	Black	Hispanic/Asian	Missing	White	Black	Hispanic/Asian	Missing	White	Black	Hispanic/Asian	Missing
PUMS	43,818	148,043	13,410	7,360	49,538	161,868	13,934	15,146	232	3,153	0	4,488
Sample	40,180	110,013	7,322	7,360	89,654	166,351	17,093	15,146	232	3,153	0	4,488

V6	Gender/Education											
	Male				Female				Missing			
	<= High School	College	Graduate School	Missing	<= High School	College	Graduate School	Missing	<= High School	College	Graduate School	Missing
PUMS	98,300	902,90	20,660	3,448	87,034	116,281	28,416	8,687	0	1,180	400	6,293
Sample	35,433	85,310	40,684	3,448	54,820	160,093	64,643	8,687	0	1,180	400	6,293

V7	Race/Education											
	Hispanic			White/Asian			Black			Missing		
	<= High School	> High School	Missing	<= High School	> High School	Missing	<= High School	> High School	Missing	<= High School	> High School	Missing
PUMS	53,059	26,427	2,903	116,816	202,041	5,417	11,338	15,995	0	3,486	13,401	10,108
Sample	41,793	74,712	2,903	36,938	247,818	5,417	8,035	16,379	0	3,486	13,401	10,108

	Max1	Max2	Max3	Max4	Max5	Max5	Max6	Max7
Max Percent Difference	182.7%	46.1%	53.3%	74.1%	133.2%	81.0%	127.5%	182.7%

Table A65. Springfield 2019: Sample and Aligned Population Weights by Step and Iteration for V5: Age x Gender

V2	Age/Gender														
	18-34			35-49			50-64			65+			Missing		
	Male	Female	Missing	Male	Female	Missing	Male	Female	Missing	Male	Female	Missing	Male	Female	Missing
PUMS	62,283	63,259	976	46,188	50,134	0	54,569	57,588	0	39,700	51,926	0	9,874	17,597	68,97
Sample	36,208	70,089	976	33,639	76,852	0	39,990	72,083	0	45,164	51,623	0	9,874	17,597	68,97
Iteration 1															
V1	33,402	66,300	976	347,60	72,987	0	41,388	74045	0	48,667	55,013	0	9,180	17,072	7,202
V2	57,456	59,839	976	47,727	47,612	0	56,476	59,156	0	42,780	55,335	0	9,180	17,072	7,202
V3	60,496	64,115	1,058	47,656	48,389	0	54,357	58,166	0	40,132	52,254	0	9,874	17,597	6,897
V4	65,361	60,132	976	46,287	49,878	0	52,756	59,360	0	37,174	54,700	0	9,874	17,597	6,898
V5	64,793	59,565	976	46,417	47,027	0	53,634	60,461	0	38,639	56,340	0	9,149	17,093	6,898
V6	65,559	59,972	988	45,679	46,882	0	53,649	60,434	0	38,732	55,825	0	9,079	17,306	6,886
V7	63,359	60,853	1,151	44,833	45,641	0	55,120	60,526	0	39,767	56,188	0	8,826	17,958	6,770
Iteration 2															
V1	63,230	60,726	1,151	45,199	46,045	0	54,489	61,368	0	39272	55,666	0	8,753	18,367	67,,26
V2	62,283	63,259	976	46,188	50,134	0	54,569	57,588	0	39,700	51,926	0	9,874	17,597	6897
V3	62,364	62,446	859	46,285	49,759	0	55,233	57,290	0	39,976	52,410	0	10,292	17,518	6,558
V4	62,728	62,881	860	46,439	49,726	0	54,892	57,223	0	39,782	52,091	0	10,422	17,475	6,471
V5	62,128	63,172	933	46,027	50,004	0	54,588	57,457	0	39,633	52,247	0	10,256	17,606	6,941
V6	62,257	63,137	910	46,138	49,999	0	54,574	57,392	0	39,688	52,185	0	10,040	17,706	6,964
V7	62,288	63,219	928	46,123	50,026	0	54,649	57,298	0	39,650	52,092	0	10,085	17,688	6,943
Iteration 3															
V1	62,208	63,193	928	46,163	50,016	0	54,626	57,389	0	39,659	52,110	0	10,055	17,700	6,944
V2	62,283	63,259	976	46,188	50,134	0	54,569	57,588	0	39,700	51,926	0	9,874	17,597	6,897
V3	62,029	62,713	928	46,140	49,904	0	54,880	57,643	0	40,006	52,380	0	9,941	17,586	6,841
V4	62,367	63,152	950	46,216	49,950	0	54,700	57,416	0	39,789	52,084	0	9,971	17,609	6,788
V5	62,168	63,222	945	46,116	50,009	0	54,648	57,504	0	39,755	52,169	0	9,945	17,581	6,929
V6	62,214	63,210	928	46,178	49,976	0	54,652	57,468	0	39,792	52,128	0	9,863	17,637	6,946
V7	62,313	63,299	935	46,147	50,035	0	54,636	57,401	0	39,725	52,028	0	9,918	17,640	6,915

The Sample row of Table A65 illustrates how close the raked marginal total weights are for the sample, compared to the estimated 2018 PUMS population total for Springfield.

Table A66 describes the maximum percent difference in marginal totals for the raked weights versus the population by raking variable and iteration. The largest percent difference in weights is for the raking variable race. After 14 iterations, the maximum percent difference was minimized, with the maximum percent difference in cell weights between the PUMs totals and the Weighted sample less than 1.177%.

Table A66. Springfield 2019: Percent Difference between Sample and Aligned Population Weights by Raking Iteration

Iteration	Max Difference	Race Max1	Age x Sex Max2	Age x RaceV1 Max3	Age x Edu Max4	Sex x RaceV1 Max5	Sex x Edu Max6	Rav2 x EduV1 Max7
1	42.21%	42.2%	17.9%	38.2%	19.7%	13.4%	13.2%	0.0%
2	9.598%	3.1%	5.0%	9.6%	8.5%	5.2%	3.5%	0.0%
3	5.576%	0.8%	4.2%	2.0%	3.4%	5.6%	2.8%	0.0%
4	4.479%	0.8%	3.9%	1.6%	1.8%	4.5%	2.1%	0.0%
5	3.457%	0.8%	2.9%	1.2%	1.0%	3.5%	1.5%	0.0%
6	2.670%	0.8%	1.9%	1.2%	0.7%	2.7%	1.2%	0.0%
7	2.105%	0.8%	1.2%	1.2%	0.4%	2.1%	0.9%	0.0%
8	1.714%	0.8%	0.8%	1.2%	0.4%	1.7%	0.7%	0.0%
9	1.447%	0.8%	0.5%	1.2%	0.4%	1.4%	0.6%	0.0%
10	1.267%	0.8%	0.3%	1.2%	0.4%	1.3%	0.6%	0.0%
11	1.177%	0.8%	0.3%	1.2%	0.4%	1.1%	0.5%	0.0%
12	1.177%	0.8%	0.3%	1.2%	0.4%	1.1%	0.5%	0.0%
13	1.177%	0.8%	0.3%	1.2%	0.4%	1.0%	0.4%	0.0%
14	1.177%	0.8%	0.3%	1.2%	0.4%	1.0%	0.4%	0.0%

A full list of the PUMS and sample respondent weight totals for the raking variables after 14 iterations is given in Table A67.

Table A67. Springfield 2019 Targeted Survey: Sample and Aligned Population Weights After 14 Iterations of Raking

V1	Race				
	Hispanic	White	Black	Asian	Missing
PUMS	81,720	313,020	27,343	11,913	26,995
Sample	82,389	312,439	27,332	11,836	26,995

V2	Age/Gender														
	18-34			35-49			50-64			65+			Missing		
	Male	Female	Missing	Male	Female	Missing	Male	Female	Missing	Male	Female	Missing	Male	Female	Missing
PUMS	62,283	63,259	976	46,188	50,134	0	54,569	57,588	0	39,700	51,926	0	9,874	17,597	6,897
Sample	62,344	63,259	976	46,086	50,081	0	54,550	57,524	0	39,761	52,008	0	9,901	17,609	68,93

V3	Age/Race											
	18-34				35-49				50-64			
	White	Black/Asian	Hispanic	Missing	White	Black/Asian	Hispanic	Missing	White	Black/Asian	Hispanic	Missing
PUMS	32,163	76,317	12,483	4,705	20,601	62,815	9,602	3,026	13,659	83,966	9,476	5,423
Sample	32,542	76,754	12,548	4,735	20,729	62,827	9,582	3,029	13,677	83,576	9,419	5,402

V3 (cont.)	Age/Race							
	65+				Missing			
	White	Black/Asian	Hispanic	Missing	White	Black/Asian	Hispanic	Missing
PUMS	7,051	77,375	5,748	2,212	8,353	12,468	1,918	11,629
Sample	7,043	76,820	5,709	2,198	8,399	12,463	1,911	11,631

V4	Age/Race											
	18-34				35-49				50-64			
	<= High School	College	Graduate School	Missing	<= High School	College	Graduate School	Missing	<= High School	College	Graduate School	Missing
PUMS	48,219	67,612	7,073	3,566	38,657	42,613	13,719	1,177	46,514	51,754	12,120	17,27
Sample	48,134	67,815	7,059	3,571	38,578	42,725	13,686	1,178	46,369	51,880	12,100	1,726

V4 (cont.)	Age/Race							
	65+				Missing			
	<= High School	College	Graduate School	Missing	<= High School	College	Graduate School	Missing
PUMS	45,617	32,971	12,344	941	6,155	12,789	4,405	11,019
Sample	45,464	33,046	12,320	940	6,154	12,837	4,398	11,014

V5	Gender/Race											
	Male				Female				Missing			
	White	Black	Hispanic/Asian	Missing	White	Black	Hispanic/Asian	Missing	White	Black	Hispanic/Asian	Missing
PUMS	43,818	148,043	13,410	7,360	49,538	161,868	13,934	15,146	232	3,153	0	4,488
Sample	44,130	147,752	13,402	7,357	49,860	161,540	13,931	15,151	234	3,147	0	4,488

V6	Gender/Education											
	Male				Female				Missing			
	<= High School	College	Graduate School	Missing	<= High School	College	Graduate School	Missing	<= High School	College	Graduate School	Missing
PUMS	98,300	90,290	20,660	3,448	87,034	116,281	28,416	8,687	0	1,180	400	6,293
Sample	97,975	90,508	20,696	3,463	86,724	116,611	28,465	86,81	0	1,182	401	6,285

V7	Race/Education											
	Hispanic			White/Asian			Black			Missing		
	<= High School	> High School	Missing	<= High School	> High School	Missing	<= High School	> High School	Missing	<= High School	> High School	Missing
PUMS	53,059	26,427	2,903	116,816	202,041	5,417	11,338	15,995	0	3,486	13,401	10,108
Sample	53,059	26,427	2,903	116,816	202,041	5,417	11,338	15,995	0	3,486	13,401	10,108

Raked Weight Alignment with Raking Variables

We summarize the alignment of raking weights with the original weights and the PUMS weight for Springfield in Tables A68-A74.

Table A68. PUMS and Sample Weight (ST_WT4) for Springfield: Race V1

	Race					
	Hispanic	White/Other	Black	Asian	Missing	All
Springfield ST-WT4	119,409	279,517	24,415	10,657	26,995	460,992
2018 PUMS	86,803	332,490	29,044	12,654	0	460,991
Raked WT	82,389	312,439	27,332	11,836	26,995	460,991

Table A69. PUMS and Sample Weight (ST_WT4) for Springfield: Age x Gender V2

		Age					
		18-34	35-49	50-64	65+	Missing	All
Springfield ST-WT4	Male	36,208	33,639	39,990	45,164	9,874	164,875
	Female	70,089	76,852	72,083	51,623	17,597	288,244
	Missing	976	0	0	0	6,897	7,874
	All	107,274	110,491	112,073	96,787	34,368	460,992
2018 PUMS	Male	66,793	49,173	58,096	42,266	0	216,328
	Female	69,796	54,913	63,078	56,876	0	244,663
	Missing	0	0	0	0	0	0
	All	136,589	104,086	121,174	99,142	0	460,991
Raked WT	Male	62,344	46,085	54,550	39,761	9,900	212,640
	Female	63,259	50,082	57,524	52,008	17,609	240,482
	Missing	976	0	0	0	6,893	7,869
	All	126,579	96,166	112,074	91,769	34,402	460,991

Table A70. PUMS and Sample Weight (ST_WT4) for Springfield: Age x Collapsed Race V3

		Age					
		18-34	35-49	50-64	65+	Missing	All
Springfield ST-WT4	Hispanic/Asian	45,905	40,686	23,437	10,733	9,304	130,065
	White/Other	48,191	61,938	77,901	79,020	12,468	279,517
	Black	8,472	4,842	5,313	4,821	967	24,415
	Missing	4,705	3,026	5,423	2,212	11,629	26,995
	All	107,274	110,491	112,073	96,787	34,368	460,992
2018 PUMS	Hispanic/Asian	42,550	27,463	19,597	9,847	0	99,457
	White/Other	84,592	69,199	94,121	84,578	0	332,490
	Black	9,447	7,424	7,456	4,717	0	29,044
	Missing	0	0	0	0	0	0
	All	136,589	104,086	121,174	99,142	0	460,991
Raked WT	Hispanic/Asian	35,025	25,451	16,762	7,438	9,549	94,225
	White/Other	76,754	62,827	83,577	76,820	12,461	312,439
	Black	10,065	4,860	6,334	5,313	760	27,332
	Missing	4,736	3,029	5,401	2,198	11,631	26,995
	All	126,579	96,166	112,074	91,769	34,402	460,991

Table A71. PUMS and Sample Weight (ST_WT4) for Springfield: Age x Education V4

		Age					
		18-34	**35-49**	**50-64**	**65+**	**Missing**	**All**
Springfield ST-WT4	<= High School	23,440	15,455	26,066	19,138	6,155	90,253
	College	66,422	62,671	56,783	47,917	12,789	246,583
	Graduate School	13,847	31,188	27,496	28,791	4,405	105,727
	Missing	3,566	1,177	1,727	941	11,019	18,429
	All	107,274	110,491	112,073	96,787	34,368	460,992
2018 PUMS	<= High School	52,551	41,450	50,035	48,814	0	192,850
	College	75,867	47,043	57,319	36,326	0	216,555
	Graduate School	8,171	15,593	13,820	14,002	0	51,586
	Missing	0	0	0	0	0	0
	All	136,589	104,086	121,174	99,142	0	460,991
Raked WT	<= High School	48,134	38,578	46,369	45,464	6,154	184,699
	College	67,815	42,725	51,879	33,045	12,837	208,301
	Graduate School	7,059	13,686	12,100	12,320	4,398	49,562
	Missing	3,571	1,178	1,726	940	11,014	18,429
	All	126,579	96,166	112,074	91,769	34,402	460,991

Table A72. PUMS and Sample Weight (ST_WT4) for Springfield: Gender x Collapsed Race V5

		Gender			
		Male	**Female**	**Missing**	**All**
Springfield ST-WT4	Hispanic/Asian	40,180	89,654	232	130,065
	White/Other	110,013	166,351	3,153	279,517
	Black	7,322	17,093	0	24,415
	Missing	7,360	15,146	4,488	26,995
	All	164,875	288,244	7,874	460,992
2018 PUMS	Hispanic/Asian	45,944	53,513	0	99,457
	White/Other	156,357	176,133	0	332,490
	Black	14,027	15,017	0	29,044
	Missing	0	0	0	0
	All	216,328	244,663	0	460,991
Raked WT	Hispanic/Asian	44,130	49,861	234	94,225
	White/Other	147,750	161,541	3,148	312,439
	Black	13,402	13,931	0	27,332
	Missing	7,358	15,150	4,487	26,995
	All	212,640	240,482	7,869	460,991

Table A73. PUMS and Sample Weight (ST_WT4) for Springfield: Gender x Education V6

		Gender			
		Male	Female	Missing	All
Springfield ST-WT4	<= High School	35,433	54,820	0	90,253
	College	85,310	160,093	1,180	246,583
	Graduate School	40,684	64,643	400	105,727
	Missing	3,448	8,687	6,293	18,429
	All	164,875	288,244	7,874	460,992
2018 PUMS	<= High School	101,303	91,547	0	192,850
	College	93,565	122,990	0	216,555
	Graduate School	21,460	30,126	0	51,586
	Missing	0	0	0	0
	All	216,328	244,663	0	460,991
Raked WT	<= High School	97,975	86,724	0	184,699
	College	90,508	116,611	1,182	208,301
	Graduate School	20,696	28,465	401	49,562
	Missing	3,462	8,681	6,286	18,429
	All	212,640	240,482	7,869	460,991

Table A74. PUMS and Sample Weight (ST_WT4) for Springfield: Collapsed Race x Education V7

		Race				
		Hispanic	White/Other	Black/Asian	Missing	All
Springfield ST-WT4	<= High School	41,793	36,138	8,836	3,486	90,253
	> High School	74,712	238,913	25,284	13,401	352,310
	Missing	2,903	4,466	951	10,108	18,429
	All	119,409	279,517	35,071	26,995	460,992
2018 PUMS	<= High School	57,289	119,018	16,543	0	192,850
	> High School	29,514	213,472	25,155	0	268,141
	Missing	0	0	0	0	0
	All	86,803	332,490	41,698	0	460,991
Raked WT	<= High School	53,059	114,720	13,434	3,486	184,699
	> High School	26,427	193,452	24,584	13,401	257,864
	Missing	2,903	4,266	1,151	10,108	18,429
	All	82,389	312,439	39,168	26,995	460,991

Step 6 Trimming Raked Weights

Weights developed via this process can have a broad range for individual respondents. While the weights provide in theory a way of obtaining unbiased estimates of population parameters, variability in weights will inflate the variance of the estimates. For this reason, trimming the weights can be desirable to improve the overall estimation accuracy. Trimming was examined in the BGPS weight development. We apply the strategy for trimming weights developed in the BGPS to weights developed for Springfield.

Trimming Raked Weights

We describe the procedure for trimming raked weights next. Let $w_{\min}$ represent the minimum weight, w_{mean} represent the mean weight, and $w_{\max}$ represent the maximum weight. We define the trimmed weight by setting the minimum and maximum weight to be a simple multiplier, m, times the average weight, w_{mean}. The initial trimmed weight is given by

$$w_{j,m}^{0} = \begin{cases} w_{\max,m} & \text{if } w_j \geq w_{\max,m} \\ w_i & \\ w_{\min,m} & \text{if } w_j \leq w_{\min,m} \end{cases} .$$

where $w_{\max,m} = m(w_{mean})$ and $w_{\min,m} = (w_{mean})/m$. By changing the minimum and maximum weight, the total weight is changed. In order to insure that the total weight is equal to the total population size, we adjust the initial trimmed weight by a factor $\frac{T}{T_m}$, where $T = \sum_{j=1}^{n} w_j$ represents the total raked weight prior to trimming, and $T_m = \sum_{j=1}^{n} w_{j,m}^{0}$ represents the total weight after trimming weights to a multiple of the mean weight. The final step in forming the trimmed weight is to multiply the initial trimmed weight by $\frac{T}{T_m}$, to form the trimmed weight

$$w_{j,m} = \left(\frac{T}{T_m}\right) w_{j,m}^{0}.$$

Determining the Extent of Trimming

In the BGPS, we determined the multiplier used to trim weights by evaluating the accuracy of estimators for values of $m = 2,3,4,5,6,7,8,9$ for the variables defined as a) problem gambler; b) at risk gambler; c) tobacco user; and d) participant in extreme sports. An unbiased estimator of the variable was assumed to be the estimator without trimming. Using this process, we found that the most accurate estimator will occur when $m = 8$. We use this multiplier to form trimmed weights.

Trimming Raked Weights for Springfield

By setting $m = 8$, the minimum and maximum raked weights are given by 50.81 and 3252.14, respectively. This results in trimming weights for 21 respondents (Table A75).

Table A75. List of Weights That are Trimmed for Springfield 2019 Targeted Survey

	Age	Sex	Race	Edu	Springfield 2019 Target Household Size Weight	Initial Raked Weight (not trimmed)	Trimmed Raked Weight
Observation	AGE_PS	SEX_PS	RACE-PS	EDU_PS	ST-WT4	ST-WT5	WT6a
1	65+	Female	Hispanic	College	113.15	26.93	50.81
2	35-49	Female	Hispanic	College	113.15	26.99	50.81
3	35-49	Female	Hispanic	College	113.15	26.99	50.81
4	50-64	Female	Hispanic	Graduate School	226.29	30.60	50.81
5	65+	Male	Hispanic	Graduate School	248.24	31.20	50.81
6	35-49	Female	Hispanic	Graduate School	226.29	32.17	50.81
7	35-49	Female	Hispanic	College	134.92	32.19	50.81
8	18-34	Female	Black	Graduate School	113.15	36.92	50.81
9	50-64	Female	Hispanic	College	134.92	41.58	50.81
10	18-34	Female	Hispanic	College	113.15	42.48	50.81
11	18-34	Female	Hispanic	College	113.15	42.48	50.81
12	35-49	Female	Hispanic	College	200.16	47.75	50.81
13	35-49	Female	Hispanic	College	200.16	47.75	50.81
14	18-34	Female	Hispanic	College	130.54	49.02	50.81
15	18-34	Female	Hispanic	College	134.92	50.66	50.81
16	18-34	Female	Hispanic	College	134.92	50.66	50.81
17	18-34	Male	White/Other	<= High School	600.49	3290.80	3252.14
18	18-34	Male	White/Other	<= High School	600.49	3290.80	3252.14
19	18-34	Male	White/Other	<= High School	600.49	3290.80	3252.14
20	35-49	Female	White/Other	<= High School	1064.25	3539.97	3252.14
21	35-49	Male	White/Other	<= High School	496.47	3621.28	3252.14

For comparison, the minimum and maximum raked weights in the Baseline Springfield Targeted Survey are given by 50.26 and 3216.78, respectively. In the Baseline Springfield Targeted Survey, this results in trimming weights for 8 respondents (see Table A32 from Baseline Springfield Targeted Survey Weighting report.

The total weight prior to trimming is 460,991, while the total weight after trimming (but prior to adjusting) is 460,414.71. We define the raked trimmed weight by multiplying the trimmed weights by the factor $\frac{T}{T_m} = \frac{460,991}{460,414.71}$. The resulting weight, ST_WT6, is the final weight for the Springfield 2019 Targeted Survey.

Appendix A3: Item Response

Table A76. Item Response Rate for Baseline and Follow-up Targeted Population Survey: Springfield

	Baseline (2015)			Follow-up (2019)		
	WEB	SAQ	Phone	WEB	SAQ	Phone
d2_R RECODED: Are you male or female?	99.4	99.6	100.0	99.1	96.4	100.0
Age (based on 2019-year of birth)	94.3	96.0	96.4	92.6	93.4	93.3
c1_RBC RECODED AND BACKCODED: Which of the following is your preferred recreational activity? Would you say…?	100.0	97.9	99.1	99.9	98.7	100.0
c2_R RECODED: Do you enjoy participating in extreme sports such as hang gliding or sky diving?	99.8	99.2	99.1	100.0	99.7	100.0
c2a_R RECODED: Do you have an internet connection either at home or at work?				98.7	97.0	93.3
c2b_R RECODED: Overall, how often do you use the Internet?				98.4	96.0	93.3
C3_R RECODED: Over the past 12 months, would you say that in general your health has been…?	100.0	99.4	98.2	100.0	99.7	100.0
C4_R RECODED: In the past 12 months, how would you rate your overall level of stress?	99.8	99.4	99.1	98.9	100.0	100.0
C5_R RECODED: In the past 12 months, how would you rate your overall level of happiness?	99.4	99.0	100.0	99.0	100.0	100.0
C6A_R RECODED: Have you smoked at least 100 cigarettes in your entire life?	100.0	99.4	100.0	100.0	99.7	100.0
C6B_R RECODED: Would you say you now smoke cigarettes…	100.0	97.1	100.0	99.9	97.7	100.0
C6C_R RECODED: Do you currently smoke cigars, pipe tobacco, or hookah tobacco (shisha), or use dipping tobacco (including snus), chewing tobacco, or snuff…?	99.8	98.7	100.0	100.0	100.0	100.0
C6D_R RECODED: During the past 30 days, how many days would you estimate you have used any form of tobacco?	98.5	91.2	100.0	96.6	88.8	100.0
C7A_R RECODED: Have you used alcohol in the past 12 months?	99.8	99.6	100.0	100.0	99.7	100.0
Alcohol use (3 categories)	99.8	99.6	100.0	100.0	99.7	100.0
C7C_R RECODED: One drink is equivalent to a 12-ounce beer, a 5-ounce glass of wine, or a drink with one shot of liquor. During the past 30 days, on the days when you drank, about how many drinks did y	95.7	98.5	93.8	94.7	70.6	86.7
Binge drinker (past 30 days)	95.4	98.5	93.8	94.5	93.1	86.7
C8_R RECODED: In the past 12 months have you used any marijuana, hallucinogens (such as LSD, mushrooms, or PCP), cocaine, heroin or opium, or any other drugs not intended for medical use?	99.6	99.8	100.0	100.0	98.7	100.0
C9A_R RECODED: Have you had any problems with drugs or alcohol in the past 12 months? By this we mean difficulties in controlling their use that have led to negative consequences for you or other peop	99.8	99.2	100.0	100.0	97.0	100.0
C9B_R RECODED: During the past 12 months, have you sought help for your use of alcohol or drugs?	99.8	99.2	100.0	100.0	97.0	100.0
C10A_R RECODED: Have you had problems with other behavior in the past 12 months such as overeating, sex or pornography, shopping, exercise, Internet chat lines, or other things?	99.8	99.8	100.0	100.0	98.0	100.0
C11A_R RECODED: In the past 30 days, have you had any serious problems with depression, anxiety or other mental health problems?	99.3	99.8	100.0	99.6	98.0	100.0
C11B_R RECODED: How about in the last 12 months?	98.9	90.8	100.0	99.5	90.4	100.0
C11D_R RECODED: During the past 12 months, did you ever seriously consider attempting suicide?	99.1	99.2	100.0	99.6	97.4	100.0
C11E_R RECODED: During the past 12 months, did you actually attempt suicide?	99.1	99.0	100.0	99.5	97.4	100.0

	Baseline (2015)			Follow-up (2019)		
	WEB	SAQ	Phone	WEB	SAQ	Phone
C12_R RECODED: Do you now have any health problem that requires you to use special equipment, such as a cane, a wheelchair, a special bed, or a special telephone?	99.4	99.4	100.0	99.9	97.7	100.0
C13_R RECODED: How would you describe your childhood?	100.0	99.2	99.1	99.9	99.0	100.0
GA1_R RECODED: Which best describes your belief about the benefit or harm that gambling has for society?	99.8	98.7	94.6	99.0	94.4	93.3
GA2_R RECODED: Do you believe that gambling is morally wrong?	99.6	98.7	100.0	99.5	97.4	100.0
GA3A_R RECODED: Which of the following best describes your opinion about legalized gambling?	99.8	96.7	95.5	98.7	94.7	86.7
GA4_R RECODED: Which of the following best describes your opinion about gambling opportunities in Massachusetts?	99.8	97.5	95.5	98.8	94.7	100.0
GA5_R RECODED: There have been 2 new casinos and one slot parlor built in Massachusetts in the past few years. What sort of overall impact do you believe these have had?	100.0	99.4	97.3	99.3	97.4	86.7
GA6A_RBC RECODED AND BACKCODED: What do you believe will be the single most positive impact for Massachusetts? Would you say…	99.6	99.2	100.0	99.4	98.0	93.3
GA6B_RBC RECODED AND BACKCODED: What do you believe will be the single most negative impact for Massachusetts? Would you say…	99.6	99.0	96.4	99.3	97.0	93.3
GA7_R RECODED: What sort of overall impact do you believe a new casino or slot parlor would have for your own community?	99.8	98.7	98.2	99.0	98.0	100.0
GY1A_R RECODED: In the past 12 months, how often have you purchased lottery tickets such as Megabucks, Powerball, or Lucky for Life?	99.8	99.6	100.0	100.0	98.7	100.0
GY2A_R RECODED: In the past 12 months, how often have you purchased instant tickets or pull tabs?	99.6	99.4	100.0	100.0	98.7	100.0
GY2C_R RECODED: In the past 12 months, how often have you purchased raffle tickets?	100.0	98.1	100.0	99.5	98.3	100.0
GY3A_R RECODED: In the past 12 months, how often have you purchased daily lottery games such as Mass Cash, All or Nothing, or Numbers Game?	99.3	99.6	100.0	99.5	97.7	100.0
GY4A_R RECODED: In the past 12 months, how often have you bet money on sporting events (this includes sports pools)?	98.9	99.4	100.0	99.4	98.0	100.0
GY5A_R RECODED: In the past 12 months, how often have you gone to a bingo hall to gamble?	99.4	98.5	100.0	99.5	98.7	100.0
GY8A_R RECODED: In the past 12 months, how many times have you gambled at a casino, racino, or slots parlor outside of Massachusetts?	99.4	88.3	100.0	99.8	87.1	100.0
gy8d_rbc - RECODED and BACKCODED:Please Specify the State	100.0	88.7	100.0	99.4	87.5	100.0
GY8E_Rbc RECODED and BACKCODED: Which specific casino, racino, or slots parlor do you most often go to? (CATI)	99.8	87.9	100.0	99.5	87.1	100.0
GY9A_R RECODED: In the past 12 months, how often have you bet on a horse race at either a horse race track or an off-track site?	99.3	99.6	100.0	99.6	100.0	100.0
gy9C_RBC- RECODED and BACKCODED:Please specify where you go most often?	100.0	97.9	100.0	100.0	100.0	100.0
GY10A_R RECODED: In the past 12 months, how often have you gambled or bet money against other people on things such as card games; golf, pool, darts, bowling; video games; board games, or poker outsid	99.6	99.4	100.0	99.5	99.7	100.0
GY11A_R RECODED: In the past 12 months, how often did you purchase high risk stocks, options or futures or day trade on the stock market?	99.1	99.6	100.0	99.4	100.0	100.0
GY12A_R RECODED: In the past 12 months, have you gambled online?	99.4	99.0	100.0	99.5	99.3	100.0
GR1_R RECODED: How important is gambling to you as a recreational activity?	98.7	99.6	100.0	100.0	92.1	100.0
GR2A_R RECODED: Has gambling replaced other recreational activities for you in the past 5 years?	98.3	98.5	100.0	99.5	93.1	100.0

	Baseline (2015)			Follow-up (2019)		
	WEB	SAQ	Phone	WEB	SAQ	Phone
PA1_R RECODED: In the past 12 months have you seen or heard any media campaigns to prevent problem gambling in Massachusetts?	98.7	98.7	99.1	99.1	98.7	100.0
PA2A_R RECODED: In the past 12 months have you been aware of any programs to prevent problem gambling (other than media campaigns) offered at your school, your place of work, in your community or else	98.9	99.8	100.0	99.3	98.7	100.0
PA2B_R RECODED: Did you participate in any of the problem gambling prevention programs that you heard of in the past 12 months?	99.3	100.0	99.1	100.0	100.0	100.0
PA3_R RECODED: Did any of these media campaigns or programs cause you to alter your own gambling behavior?	98.7	99.0	96.4	99.6	99.0	100.0
GPO1_R RECODED: What portion of your close friends and family members are regular gamblers?	98.9	99.4	98.2	99.3	97.7	93.3
GPO2_R RECODED: During the last 12 months, has there been a person in your life that you consider gambles too much?	99.3	99.6	100.0	99.5	98.3	93.3
GPO3_RBC- RECODED and BACKCODED:Please specify this persons relationship to you.	99.1	98.3	99.1	99.3	95.0	93.3
GPO4_1_Rbc RECODED and backcoded: In what ways has this persons gambling affected you during the last 12 months? Reduced time spent socializing? (CATI)	99.3	81.6	99.1	99.5	98.3	93.3
GPO5_R RECODED: Overall, on a scale from 1 to 10 how much has this person's gambling affected you negatively during the last 12 months?	98.9	98.5	100.0	99.4	94.7	93.3
GP1_R RECODED: Thinking about the past 12 months, have you bet more than you could really afford to lose?	99.1	99.4	100.0	99.9	93.7	100.0
GP2_R RECODED: Thinking about the past 12 months, have you felt guilty about the way you gamble or what happens when you gamble?	99.1	99.0	100.0	99.5	92.7	100.0
GP3_R RECODED: In the past 12 months, have you needed to gamble with larger amounts of money to get the same feeling of excitement?	98.2	99.2	100.0	99.3	92.7	100.0
GP4_R RECODED: In the past 12 months, when you gambled, did you go back another day to try to win back the money you lost?	98.2	99.2	100.0	99.3	92.7	100.0
GP5A_R RECODED: In the past 12 months, have you borrowed money or sold anything to get money to gamble?	98.7	99.4	100.0	99.3	92.7	100.0
GP5B_R RECODED: In the past 12 months, about how much money have you borrowed or obtained from selling possessions in order to gamble?	98.5	99.4	100.0	100.0	95.0	100.0
GP6A_R RECODED: In the past 12 months, has your gambling caused any financial problems for you or your household?	98.7	99.0	100.0	99.4	92.7	100.0
GP6B_R RECODED: In the past 12 months, have you filed for bankruptcy because of gambling?	98.7	99.6	100.0	99.9	95.0	100.0
GP7A_R RECODED: In the past 12 months, has your gambling caused you any health problems, including stress or anxiety?	98.3	99.2	100.0	99.4	92.7	100.0
GP7B_R RECODED: In the past 12 months have these health problems caused you to seek medical or psychological help?	98.7	99.6	100.0	99.9	95.0	100.0
GP8_R RECODED: In the past 12 months, have people criticized your betting or told you that you had a gambling problem, regardless of whether or not you thought it was true?	98.3	99.0	100.0	99.5	94.1	100.0
GP9_R RECODED: In the past 12 months, have you felt that you might have a problem with gambling?	98.0	99.2	100.0	99.5	94.1	100.0
GP10A_R RECODED: Has your involvement in gambling caused significant mental stress in the form of guilt, anxiety, or depression for you or someone close to you in the past 12 months?	98.3	99.0	100.0	99.5	92.7	100.0
GP10B_R RECODED: In the past 12 months, have you thought of committing suicide because of gambling?	98.5	99.6	99.1	100.0	95.0	100.0
GP10C_R RECODED: In the past 12 months, have you attempted suicide because of gambling?	98.5	99.6	99.1	100.0	95.0	100.0
GP10D_R RECODED: Would you like to know about the free gambling and mental health treatment services in your local area?	98.5	99.6	99.1	100.0	95.0	100.0

	Baseline (2015)			Follow-up (2019)		
	WEB	SAQ	Phone	WEB	SAQ	Phone
GP11A_R RECODED: Has your involvement in gambling caused significant problems in your relationship with your spouse/partner or important friends or family in the past 12 months?	98.3	99.2	100.0	99.1	94.1	100.0
GP11B_R RECODED: In the past 12 months, has your involvement in gambling caused an instance of domestic violence in your household?	98.5	99.6	100.0	100.0	95.0	100.0
GP11C_R RECODED: In the past 12 months, has your involvement in gambling resulted in separation or divorce?	98.5	99.6	100.0	100.0	94.7	100.0
GP12A_R RECODED: In the past 12 months, has your involvement in gambling caused you to repeatedly neglect your children or family?	98.5	99.2	100.0	99.3	93.7	100.0
GP12B_R RECODED: In the past 12 months, has child welfare services become involved because of your gambling?	98.5	99.6	100.0	99.9	95.0	100.0
GP13A_R RECODED: Has your involvement in gambling caused significant work or school problems for you or someone close to you in the past 12 months or caused you to miss a significant amount of time of	97.8	99.0	100.0	99.6	94.1	100.0
GP13B_R RECODED: In the past 12 months, about how many work or school days have you lost due to gambling?	98.5	99.6	100.0	99.9	95.0	100.0
GP13C_R RECODED: In the past 12 months, have you lost your job or had to quit school due to gambling?	98.5	99.4	100.0	99.9	94.4	100.0
GP13D_R RECODED: In the past 12 months, did anyone in this household receive public assistance or other welfare payments as a result of losing your job because of gambling?	98.5	99.6	100.0	99.9	95.0	100.0
GP13E_R RECODED: Roughly how much money did you receive from public assistance in the past 12 months?	98.5	99.6	100.0	99.9	95.0	100.0
GP14A_R RECODED: In the past 12 months, has your involvement in gambling caused you or someone close to you to write bad checks, take money that didn't belong to you or commit other illegal acts to su	98.3	99.4	100.0	99.1	93.1	100.0
GP14B_R RECODED: In the past 12 months, about how much money have you illegally obtained in order to gamble?	98.5	99.6	99.1	99.8	95.0	100.0
GP14C_R RECODED: In the past 12 months, has your gambling been a factor in your committing a crime for which you have been arrested?	98.5	99.6	99.1	99.8	95.0	100.0
GP14D_R RECODED: Were you convicted for this crime?	98.5	99.6	99.1	99.8	95.0	100.0
GP14G_R RECODED: Were you incarcerated for this crime?	98.5	99.6	99.1	99.8	95.0	100.0
GP14H_R RECODED: For how many days were you incarcerated?	98.5	99.6	99.1	99.8	95.0	100.0
GP15_R RECODED: In the past 12 months, have you often gambled longer, with more money or more frequently than you intended to?	98.3	98.5	100.0	98.8	93.7	100.0
GP16A_R RECODED: In the past 12 months, have you made attempts to either cut down, control or stop gambling?	97.6	97.7	98.2	98.8	93.7	100.0
GP16B_R RECODED: Were you successful in these attempts to cut down, control or stop gambling?	98.5	99.6	98.2	99.8	95.0	100.0
GP17_R RECODED: In the past 12 months, is there anyone else who would say that you had difficulty controlling your gambling, regardless of whether you agreed with them or not?	98.0	98.1	100.0	99.1	93.4	100.0
GP18_R RECODED: In the past 12 months, would you say you have been preoccupied with gambling?	98.3	98.3	100.0	99.1	93.7	100.0
GP19_R RECODED: In the past 12 months, when you did try cutting down or stopping did you find you were very restless or irritable or that you had strong cravings for it?	97.4	98.1	98.2	99.1	92.7	100.0
GP20_R RECODED: In the past 12 months, did you find you needed to gamble with larger and larger amounts of money to achieve the same level of excitement?	97.8	98.7	100.0	98.9	93.7	100.0
GP21_R RECODED: Are there particular types of gambling that have contributed to your problems more than others?	98.3	98.7	100.0	99.4	90.8	100.0
GP23A_R RECODED: Have you wanted help for gambling problems in the past 12 months?	98.3	98.7	100.0	99.4	90.8	100.0

	Baseline (2015)			Follow-up (2019)		
	WEB	SAQ	Phone	WEB	SAQ	Phone
GP23B_R RECODED: Have you sought help for gambling problems in the past 12 months?	98.3	98.7	100.0	99.4	90.8	100.0
GP23D_R RECODED: How helpful was this?	98.3	98.7	100.0	99.4	90.8	100.0
GP23E_R RECODED: Have you excluded yourself from any casino or slots parlor in the past 12 months?	98.3	98.7	100.0	99.3	90.8	100.0
GP24_R RECODED: Have you had problems with gambling in your lifetime prior to the past 12 months?	98.3	98.7	100.0	99.3	90.8	100.0
Canadian Problem Gambling Index	98.3	99.0	100.0	99.6	93.1	100.0
D4_R RECODED: At present are you…?	95.9	97.9	99.1	95.1	95.7	100.0
D5_R RECODED: How many children under 18 years old live in your household?	94.3	95.4	99.1	92.9	92.7	100.0
D6_R RECODED: What is the highest degree or level of school you have completed?	97.8	98.7	100.0	96.2	96.7	100.0
Employment (6 categories)	96.9	99.2	100.0	96.0	98.0	100.0
D7B_R RECODED: Have you ever served on active duty in the U.S. Armed Forces, military Reserves, or National Guard?	97.8	99.0	100.0	96.9	97.4	100.0
D8_RBC RECODED and BACKCODED: What type of healthcare coverage do you have?	94.6	97.5	94.6	92.4	98.0	100.0
D9_RBC RECODED AND BACKCODED: Do you own the place where you currently live, pay rent or something else?	95.9	98.3	99.1	93.6	98.7	100.0
Household income (6 categories)	81.1	89.5	88.4	80.0	87.8	60.0
Current debt	84.1	91.8	83.0	84.2	90.4	46.7
D12_R RECODED: Were you born in the United States?	97.0	98.3	98.2	96.9	96.7	93.3
D12A_R RECODED: Do you live in Massachusetts for 6 or more months out of the year?	98.2	96.9	98.2	97.3	96.4	93.3
D13_R RECODED: Are you Hispanic or Latino?	95.9	97.1	98.2	94.9	96.0	86.7
ethnicity1	93.5	97.7	96.4	94.0	95.7	80.0
Current tobacco use	100.0	96.9	100.0	99.9	98.0	100.0
Education (6 categories)	97.8	98.7	100.0	96.2	96.7	100.0
Marital status (5 categories)	95.9	97.9	99.1	95.1	95.7	100.0

Appendix B: Tabular Results

Table 2. Attitudes: Targeted Springfield Survey (Weighted)

		Baseline (2015)			Follow-up (2019)			
		N	%	95% CI	N	%	95% CI	p-value
Opinions about legalized gambling	All types of gambling should be illegal	58,002	13.1	(10.4, 16.4)	56,139	12.5	(10.2, 15.3)	0.5134 MH
	Some types of gambling should be legal and some illegal	249,086	56.2	(52.0, 60.3)	235,224	52.3	(48.1, 56.5)	
	All types of gambling should be legal	136,479	30.8	(27.1, 34.7)	158,321	35.2	(31.2, 39.5)	
Beliefs about gambling availability in MA	Gambling is too widely available	90,416	20.3	(17.1, 23.9)	124,854	27.7	(24.2, 31.6)	0.0041 MH
	The current availability of gambling is	255,390	57.3	(53.1, 61.5)	242,371	53.8	(49.6, 58.0)	
	Gambling is not available enough	99,549	22.4	(18.9, 26.2)	83,097	18.5	(15.3, 22.1)	
Perceived impact of expanded gambling in MA	Harmful	175,143	38.8	(34.9, 42.8)	99,893	21.9	(18.8, 25.3)	<.0001 MH
	Neither beneficial nor harmful	76,020	16.8	(13.7, 20.5)	159,956	35.0	(31.1, 39.1)	
	Beneficial	200,649	44.4	(40.3, 48.6)	196,742	43.1	(38.9, 47.4)	
Perceived impact of expanded gambling on own community	Harmful	162,312	36.5	(32.7, 40.4)	163,863	35.9	(32.0, 40.0)	0.9393 MH
	Neither beneficial nor harmful	101,489	22.8	(19.4, 26.6)	146,124	32.0	(28.2, 36.1)	
	Beneficial	181,234	40.7	(36.7, 44.9)	146,424	32.1	(28.3, 36.1)	
Attitudes about gambling as a recreational activity	Not at all important	206,689	62.4	(57.7, 66.9)	227,981	67.9	(62.9, 72.6)	0.0005 MH
	Not very important	91,109	27.5	(23.6, 31.9)	76,048	22.7	(18.7, 27.1)	
	Somewhat important	31,010	9.4	(6.7, 12.9)	28,280	8.4	(5.5, 12.7)	
	Very important	***2,296***	***0.7***	***(0.3, 1.8)***	***3,251***	***1.0***	***(0.4, 2.5)***	
Perceived benefit or harm of gambling to society	The harm outweighs the benefits	231,300	51.5	(47.3, 55.7)	287,824	64.1	(59.9, 68.1)	0.0013 MH
	The benefits are about equal to the harm	161,340	35.9	(31.9, 40.1)	115,221	25.7	(22.1, 29.6)	
	The benefits outweigh the harm	56,565	12.6	(9.8, 16.0)	46,030	10.2	(7.8, 13.3)	
Perceived morality of gambling	No	370,550	81.8	(78.0, 85.1)	388,527	84.9	(81.7, 87.6)	0.1634 CS
	DON'T KNOW		---			---		
	Yes	81,941	18.1	(14.8, 21.9)	69,138	15.1	(12.4, 18.3)	

Weighted N is the total number of respondents who answered the question weighted to the MA population
RED indicates estimates are unreliable, relative standard error >30%
If cell size is 5 or less, results are set to dash (---)
MH indicates that Mann Whitney U/ Wilcoxon rank sum test used
CS indicates Chi-square test used

Table 3. Gambling Behaviors: Targeted Springfield Survey (Weighted)

	Baseline (2015)			Follow-up (2019)			
	N	%	95% CI	N	%	95% CI	p-value
All gambling	320,630	70.9	(66.8, 74.6)	338,235	73.8	(70.0, 77.2)	0.2884 CS
All lottery	281,177	62.0	(57.9, 66.0)	281,136	61.2	(57.1, 65.2)	0.7825 CS
Traditional lottery	260,633	57.6	(53.4, 61.6)	259,902	56.5	(52.4, 60.6)	0.7241 CS
Instant games	186,668	41.2	(37.3, 45.3)	195,381	42.4	(38.3, 46.7)	0.6818 CS
Daily games	78,172	17.3	(14.0, 21.1)	110,404	24.1	(20.6, 28.0)	0.0083 CS
Raffles	142,545	31.6	(28.1, 35.3)	141,784	31.0	(27.2, 35.1)	0.8352 CS
Casino out of state	95,489	22.5	(19.4, 25.9)	104,748	23.4	(20.1, 27.1)	0.6975 CS
Casino in MA		NA		145,631	32.1	(28.5, 36.0)	
Casino anywhere	95,489	22.5	(19.4, 25.9)	162,215	36.5	(32.6, 40.7)	0.0001 CS
Sports betting	43,858	9.7	(7.7, 12.2)	49,928	10.9	(8.2, 14.4)	0.5325 CS
Private betting	43,426	9.6	(7.3, 12.4)	38,497	8.4	(6.3, 11.1)	0.5030 CS
Horse racing	14,665	3.2	(2.2, 4.8)	13,962	3.0	(2.0, 4.7)	0.8286 CS
Bingo	18,538	4.1	(3.0, 5.6)	24,603	5.4	(3.7, 7.8)	0.3003 CS
Online	***4,555***	***1.0***	***(0.5, 2.0)***	7,000	1.5	(0.9, 2.5)	0.3165 CS

Weighted N is the total number of respondents who answered the question weighted to the MA population
RED indicates estimates are unreliable, relative standard error >30%
If cell size is 5 or less, results are set to dash (---)
CS indicates Chi-square test used

Table 4. Gambling Behaviors: Targeted Springfield Survey (Weighted) by Gender

	Males					Females					
	BTPS_SP (2015)		FTPS_SP (2019)			BTPS_SP (2015)		FTPS_SP (2019)			Overall
	%	95% CI	%	95% CI	p-value[cs]	%	95% CI	%	95% CI	p-value[cs]	p-value[CMH]
All gambling	74.9	(67.8, 80.9)	76.9	(70.8, 82.2)	0.6417	67.5	(62.7, 72.0)	71.7	(67.0, 75.9)	0.2093	0.2414
All lottery	69.0	(61.8, 75.3)	66.4	(59.6, 72.6)	0.5873	56.3	(51.5, 61.0)	57.6	(52.5, 62.4)	0.7121	0.8598
Traditional lottery	65.5	(58.4, 72.0)	61.6	(54.6, 68.1)	0.4237	50.9	(46.2, 55.6)	53.1	(48.1, 58.0)	0.5369	0.8211
Instant games	44.4	(37.6, 51.5)	48.0	(41.0, 55.0)	0.4901	38.5	(34.1, 43.0)	38.2	(33.3, 43.2)	0.9328	0.6207
Daily games	22.1	(16.2, 29.3)	25.1	(19.6, 31.6)	0.5034	13.3	(10.4, 16.9)	23.6	(19.2, 28.5)	0.0004	0.0085
Raffles	32.0	(26.1, 38.5)	31.2	(25.0, 38.2)	0.8629	31.5	(27.5, 35.8)	31.0	(26.4, 35.9)	0.8592	0.8059
Casino out of state	21.7	(16.8, 27.5)	20.3	(15.5, 26.2)	0.7298	23.3	(19.6, 27.5)	26.9	(22.5, 31.8)	0.2512	0.6009
Casino in MA	NA		30.5	(24.7, 36.9)		NA		34.4	(29.8, 39.4)		
All Casino	21.7	(16.8, 27.5)	34.0	(27.9, 40.7)	0.0042	23.3	(19.6, 27.5)	39.6	(34.6, 44.8)	<0.0001	<0.0001
Sports betting	15.1	(11.3, 19.9)	18.5	(13.1, 25.4)	0.3800	5.1	(3.6, 7.3)	4.5	(3.0, 6.6)	0.6088	0.5255
Private betting	14.0	(9.8, 19.6)	10.9	(7.2, 16.1)	0.3479	5.9	(4.1, 8.4)	6.3	(4.3, 9.1)	0.7891	0.4884
Horse racing	4.8	(2.8, 8.1)	4.4	(2.6, 7.5)	0.8550	2.0	(1.2, 3.2)	***1.9***	***(0.9, 4.0)***	0.9435	0.8463
Bingo	***1.3***	***(0.5, 2.9)***	***5.6***	***(2.8, 10.7)***	0.0311	6.6	(4.7, 9.1)	5.3	(3.7, 7.7)	0.4170	0.2734
Online	***1.7***	***(0.7, 3.8)***	***1.3***	***(0.6, 2.9)***	0.6982	---		***1.5***	***(0.7, 3.0)***	0.0729	0.4527

RED indicates estimates are unreliable, relative standard error >30%
If cell size is 5 or less, results are set to dash (---)
CS indicates Chi-square test used
CMH indicates Cochran-Mantel-Haesznal test used

Table 5. Gambling Behaviors: Targeted Springfield Survey (Weighted) by Age

	18-34					35-49				
	BTPS_SP (2015)		FTPS_SP (2019)			BTPS_SP (2015)		FTPS_SP (2019)		
	%	95% CI	%	95% CI	p-value^{CS}	%	95% CI	%	95% CI	p-value^{CS}
All gambling	68.1	(58.6, 76.3)	63.5	(54.9, 71.2)	0.4524	68.3	(57.4, 77.4)	80.0	(72.2, 86.1)	0.0693
All lottery	62.5	(52.9, 71.2)	45.8	(37.3, 54.5)	0.0101	59.2	(49.0, 68.7)	66.2	(56.9, 74.4)	0.0371
Traditional lottery	57.1	(47.5, 66.1)	43.6	(35.2, 52.4)	0.0401	53.5	(43.6, 63.1)	59.5	(49.6, 68.6)	0.3975
Instant games	41.8	(33.0, 51.2)	28.2	(20.7, 37.2)	0.0308	35.2	(27.2, 44.1)	46.3	(36.5, 56.3)	0.1128
Daily games	22.9	(15.0, 33.2)	12.4	(8.1, 18.6)	0.0573	16.5	(10.1, 25.7)	24.0	(16.2, 34.0)	0.2142
Raffles	25.2	(18.2, 34.0)	18.8	(12.9, 26.7)	0.2290	30.0	(22.9, 38.1)	32.6	(24.1, 42.5)	0.6657
Casino out of state	25.7	(18.8, 33.9)	19.1	(13.8, 25.8)	0.1808	23.3	(16.8, 31.4)	20.9	(14.6, 29.0)	0.6424
Casino in MA	NA		30.1	(23.2, 38.0)		NA		29.0	(22.0, 37.1)	
All Casino	25.7	(18.8, 33.9)	32.7	(25.6, 40.8)	0.2018	23.3	(16.8, 31.4)	34.4	(26.3, 43.6)	0.0565
Sports betting	10.5	(6.2, 17.4)	14.1	(7.9, 23.8)	0.4697	10.7	(6.9, 16.2)	15.6	(9.7, 24.2)	0.2641
Private betting	17.5	(11.2, 26.4)	11.8	(6.6, 20.3)	0.2648	6.1	(3.7, 9.9)	8.2	(4.9, 13.6)	0.4148
Horse racing	---		*4.2*	*(1.9, 8.9)*	0.2428	*2.6*	*(1.2, 5.6)*	---		0.1144
Bingo	*3.9*	*(1.9, 7.8)*	*4.4*	*(2.1, 8.7)*	0.8457	*2.4*	*(1.1, 4.9)*	*6.3*	*(3.0, 12.6)*	0.1185
Online	---		---			---		*3.0*	*(1.4, 6.7)*	0.2113

	50-64					65+					
	BTPS_SP (2015)		FTPS_SP (2019)			BTPS_SP (2015)		FTPS_SP (2019)			
	%	95% CI	%	95% CI	p-value^{CS}	%	95% CI	%	95% CI	p-value^{CS}	p-value^{CMH}
All gambling	78.1	(71.8, 83.4)	81.0	(73.7, 86.7)	0.5130	70.8	(64.8, 76.1)	75.5	(68.3, 81.5)	0.2905	0.2307
All lottery	67.2	(60.5, 73.3)	72.4	(64.8, 78.9)	0.2930	60.6	(54.3, 66.6)	65.7	(58.0, 72.6)	0.3015	0.8598
Traditional lottery	64.4	(57.6, 70.7)	66.7	(58.9, 73.7)	0.6494	56.3	(49.9, 62.6)	60.8	(53.0, 68.1)	0.3780	0.8022
Instant games	46.3	(39.3, 53.4)	56.1	(47.9, 63.9)	0.0774	41.8	(35.3, 48.4)	45.5	(37.6, 53.6)	0.4855	0.5930
Daily games	15.2	(10.9, 20.7)	30.6	(23.1, 39.3)	0.0020	13.9	(10.0, 19.0)	32.0	(24.4, 40.7)	0.0003	0.0080
Raffles	34.9	(28.5, 41.9)	40.8	(32.6, 49.6)	0.2921	36.2	(30.1, 42.8)	34.3	(27.1, 42.3)	0.7070	0.8165
Casino out of state	21.4	(16.3, 27.5)	29.5	(22.0, 38.4)	0.1123	20.5	(15.4, 26.9)	25.7	(19.1, 33.7)	0.2752	0.6527
Casino in MA	NA		40.8	(32.7, 49.4)		NA		31.9	(24.7, 40.0)		
All Casino	21.4	(16.3, 27.5)	47.0	(38.6, 55.6)	<0.0001	20.5	(15.4, 26.9)	34.9	(27.3, 43.3)	0.0048	<0.0001
Sports betting	9.7	(6.6, 14.0)	*7.6*	*(3.6, 15.4)*	0.5454	7.5	(4.8, 11.6)	6.6	(4.0, 10.7)	0.7089	0.5176
Private betting	8.2	(5.4, 12.3)	8.9	(5.5, 14.0)	0.8031	4.2	(2.6, 6.8)	4.8	(2.7, 8.4)	0.7291	0.5560
Horse racing	*5.0*	*(2.5, 9.8)*	*2.0*	*(0.8, 4.8)*	0.1304	*3.7*	*(2.0, 6.7)*	*4.0*	*(2.0, 7.9)*	0.8777	0.7695
Bingo	*2.8*	*(1.3, 6.1)*	*5.3*	*(2.0, 13.4)*	0.3885	6.9	(4.4, 10.6)	*6.8*	*(3.5, 12.8)*	0.9733	0.3209
Online	---		---			---		---			0.3095

RED indicates estimates are unreliable, relative standard error >30%; If cell size is 5 or less, results are set to dash (---)

CS indicates Chi-square test used ; CMH indicates Cochran-Mantel-Haesznal test used

Table 6. Gambling Behaviors: Targeted Springfield Survey (Weighted) by Race/Ethnicity

	Hispanic/Black/Asian					White/Other					
	BTPS_SP (2015)		FTPS_SP (2019)			BTPS_SP (2015)		FTPS_SP (2019)			Overall
	%	95% CI	%	95% CI	p-value^{CS}	%	95% CI	%	95% CI	p-value^{CS}	p-value^{CMH}
All gambling	59.4	(48.6, 69.3)	59.3	(51.7, 66.4)	0.9880	74.6	(70.3, 78.4)	81.1	(76.9, 84.6)	0.0233	0.1560
All lottery	53.2	(42.8, 63.4)	49.8	(42.5, 57.1)	0.5994	64.7	(60.3, 68.8)	66.6	(61.6, 71.3)	0.5603	0.9314
Traditional lottery	48.3	(38.2, 58.5)	47.7	(40.4, 55.0)	0.9260	60.3	(55.9, 64.6)	60.6	(55.4, 65.6)	0.9349	0.8362
Instant games	35.7	(27.2, 45.4)	34.6	(28.0, 41.9)	0.8479	42.7	(38.3, 47.3)	46.6	(41.3, 52.0)	0.2799	0.5754
Daily games	12.6	(6.7, 22.5)	25.8	(19.8, 33.0)	0.0111	19.3	(15.5, 23.8)	23.8	(19.4, 28.8)	0.1648	0.0067
Raffles	16.8	(11.1, 24.5)	20.9	(15.3, 27.9)	0.3827	35.8	(31.6, 40.1)	36.0	(31.0, 41.3)	0.9550	0.9786
Casino out of state	16.5	(10.6, 24.6)	16.4	(12.3, 21.4)	0.9786	23.8	(20.2, 27.7)	26.7	(22.2, 31.7)	0.3506	0.5693
Casino in MA	NA		29.6	(23.6, 36.5)		NA		33.7	(29.0, 38.8)		
All Casino	16.5	(10.6, 24.6)	31.6	(25.4, 38.6)	0.0026	23.8	(20.2, 27.7)	39.2	(34.1, 44.5)	<0.0001	<0.0001
Sports betting	***5.5***	***(2.6, 11.3)***	9.2	(5.3, 15.6)	0.2593	10.9	(8.5, 13.8)	11.8	(8.3, 16.5)	0.7121	0.4987
Private betting	***5.7***	***(2.7, 11.4)***	5.6	(3.1, 9.7)	0.9611	10.6	(7.9, 14.2)	10.0	(7.2, 13.8)	0.8030	0.5762
Horse racing	--		---			4.0	(2.6, 6.1)	3.6	(2.2, 5.9)	0.7494	0.8949
Bingo	***4.5***	***(2.3, 8.6)***	5.5	(3.1, 9.5)	0.6446	3.9	(2.7, 5.7)	5.6	(3.5, 9.0)	0.2741	0.2989
Online	--		***1.9***	***(0.7, 4.6)***	0.2503	***1.0***	***(0.4, 2.2)***	***1.4***	***(0.8, 2.6)***	0.4699	0.3395

RED indicates estimates are unreliable, relative standard error >30%
If cell size is 5 or less, results are set to dash (---)
CS indicates Chi-square test used
CMH indicates Cochran-Mantel-Haesznal test used
NOTE: The same race/ethnicity groups used in other SEIGMA reports are used in this report to maintain consistency. Given the small proportion of individuals who identified as “Other” (8 out of 876 or 0.09% in the BTPS and 15 out of 693 or 2.16% in the FTPS), grouping these individuals with individuals who identified as White primarily highlights differences between Blacks, Hispanics and Asians, on the one hand, and Whites, on the other.

Table 7. Gambling Behaviors: Targeted Springfield Survey (Weighted) by Education

	High School					College				
	BTPS_SP (2015)		FTPS_SP (2019)			BTPS_SP (2015)		FTPS_SP (2019)		
	%	95% CI	%	95% CI	p-value[CS]	%	95% CI	%	95% CI	p-value[CS]
All gambling	66.2	(58.0, 73.6)	72.8	(65.1, 79.3)	0.2297	73.4	(68.5, 77.7)	76.4	(72.1, 80.2)	0.3312
All lottery	61.7	(53.6, 69.2)	64.6	(56.3, 72.1)	0.6181	62.4	(57.3, 67.3)	60.6	(55.7, 65.2)	0.5979
Traditional lottery	58.9	(50.8, 66.5)	60.5	(52.0, 68.3)	0.7838	56.3	(51.1, 61.4)	55.2	(50.3, 60.0)	0.7615
Instant games	41.2	(33.7, 49.1)	51.2	(42.9, 59.6)	0.0905	41.2	(36.3, 46.3)	38.4	(33.8, 43.2)	0.4128
Daily games	16.9	(11.2, 24.6)	29.2	(22.2, 37.5)	0.0169	18.9	(14.7, 24.0)	23.8	(19.9, 28.2)	0.1246
Raffles	22.1	(16.6, 28.9)	28.2	(20.8, 37.0)	0.2501	36.5	(31.7, 41.6)	32.9	(28.6, 37.6)	0.3011
Casino out of state	18.6	(13.5, 25.0)	21.4	(15.2, 29.3)	0.5411	26.2	(21.9, 31.0)	26.6	(22.5, 31.2)	0.9061
Casino in MA	NA		27.7	(21.0, 35.6)		NA		37.9	(33.2, 42.8)	
All Casino	18.6	(13.5, 25.0)	33.6	(26.0, 42.2)	0.0037	26.2	(21.9, 31.0)	41.2	(36.4, 46.1)	<0.0001
Sports betting	5.2	(3.1, 8.6)	12.2	(6.9, 20.4)	0.0615	14.0	(10.6, 18.4)	10.7	(7.8, 14.5)	0.1975
Private betting	*5.6*	*(2.6, 11.5)*	7.9	(4.4, 13.8)	0.4492	12.7	(9.4, 17.1)	9.5	(6.7, 13.4)	0.2139
Horse racing	*2.5*	*(1.0, 6.0)*	---		0.9728	3.7	(2.2, 6.1)	3.9	(2.3, 6.5)	0.8885
Bingo	3.3	(1.9, 5.6)	*7.2*	*(3.9, 13.1)*	0.1059	5.1	(3.3, 7.7)	4.4	(2.9, 6.8)	0.6805
Online	---		*1.6*	*(0.7, 3.6)*	0.0440	*1.9*	*(0.9, 4.0)*	*1.5*	*(0.7, 3.0)*	0.6499

	Graduate School					
	BTPS_SP (2015)		FTPS_SP (2019)			
	%	95% CI	%	95% CI	p-value[CS]	p-value[CMH]
All gambling	76.7	(70.9, 81.7)	73.1	(65.3, 79.7)	0.4371	0.2256
All lottery	61.5	(54.8, 67.8)	58.7	(50.8, 66.1)	0.5821	0.8775
Traditional lottery	57.6	(50.8, 64.1)	53.9	(46.2, 61.4)	0.4777	0.8005
Instant games	40.6	(33.8, 47.7)	33.7	(26.8, 41.3)	0.1800	0.6124
Daily games	11.8	(7.9, 17.2)	10.2	(6.5, 15.5)	0.6208	0.0071
Raffles	44.9	(38.2, 51.8)	36.5	(29.7, 44.0)	0.1018	0.8842
Casino out of state	21.6	(16.1, 28.4)	22.3	(16.8, 29.0)	0.8705	0.5825
Casino in MA	NA		27.9	(21.7, 35.1)		
All Casino	21.6	(16.1, 28.4)	31.5	(25.0, 38.9)	0.0349	<0.0001
Sports betting	9.6	(6.6, 13.8)	8.4	(5.1, 13.6)	0.6764	0.5041
Private betting	11.1	(7.3, 16.4)	8.1	(4.9, 13.0)	0.3244	0.5700
Horse racing	4.1	(2.3, 7.4)	---		0.2138	0.9030
Bingo	---		*2.9*	*(1.4, 5.9)*	0.3689	0.3314
Online	---		---			0.3210

RED indicates estimates are unreliable, relative standard error >30%

If cell size is 5 or less, results are set to dash (---)

CS indicates Chi-square test used

CMH indicates Cochran-Mantel-Haesznal test used

Table 8. Gambling Behaviors: Targeted Springfield Survey (Weighted) by Income

	<$50K					$50-<100K				
	BTPS_SP (2015)		FTPS_SP (2019)			BTPS_SP (2015)		FTPS_SP (2019)		
	%	95% CI	%	95% CI	p-value[CS]	%	95% CI	%	95% CI	p-value[CS]
All gambling	61.7	(54.7, 68.2)	73.8	(67.4, 79.3)	0.0099	77.9	(70.2, 84.1)	76.4	(69.3, 82.4)	0.7629
All lottery	56.6	(49.8, 63.3)	65.7	(59.0, 71.9)	0.0607	70.9	(63.3, 77.6)	57.6	(49.9, 65.0)	0.0132
Traditional lottery	52.0	(45.2, 58.7)	61.8	(54.9, 68.2)	0.0476	64.1	(56.3, 71.3)	52.4	(44.7, 60.0)	0.0333
Instant games	37.1	(31.0, 43.6)	46.4	(39.2, 53.8)	0.0643	52.2	(44.3, 59.9)	43.4	(35.7, 51.3)	0.1187
Daily games	16.5	(11.6, 23.0)	27.2	(21.1, 34.4)	0.0178	19.5	(13.7, 27.1)	26.3	(19.5, 34.4)	0.1884
Raffles	19.9	(15.5, 25.2)	22.1	(16.0, 29.7)	0.6057	39.9	(32.6, 47.6)	37.7	(30.4, 45.6)	0.6963
Casino out of state	16.7	(12.7, 21.7)	19.4	(14.8, 25.0)	0.4459	29.6	(23.0, 37.2)	25.2	(19.2, 32.3)	0.3727
Casino in MA	NA		30.4	(24.4, 37.2)		NA		35.4	(28.4, 43.2)	
All Casino	16.7	(12.7, 21.7)	33.9	(27.4, 40.9)	<0.0001	29.6	(23.0, 37.2)	39.5	(32.1, 47.4)	0.0665
Sports betting	***3.6***	***(1.9, 6.5)***	***7.4***	***(3.9, 13.5)***	0.1429	14.5	(9.9, 20.8)	12.6	(7.7, 19.9)	0.6330
Private betting	4.0	(2.3, 7.0)	7.3	(4.7, 11.3)	0.1025	11.8	(7.5, 18.2)	8.5	(4.7, 14.8)	0.3642
Horse racing	***2.1***	***(0.9, 5.2)***	***1.7***	***(0.7, 4.3)***	0.7107	***4.7***	***(2.5, 8.8)***	***5.7***	***(3.0, 10.6)***	0.6839
Bingo	4.2	(2.7, 6.4)	5.7	(3.2, 9.7)	0.4231	***3.7***	***(1.9, 6.8)***	4.2	(2.5, 7.2)	0.7249
Online	---		***1.9***	***(0.8, 4.2)***	0.0385	***1.8***	***(0.7, 4.5)***	---		0.5390

	$100K +					
	BTPS_SP (2015)		FTPS_SP (2019)			
	%	95% CI	%	95% CI	p-value[CS]	p-value[CMH]
All gambling	85.2	(79.6, 89.5)	82.6	(75.5, 88.0)	0.5239	0.3294
All lottery	66.0	(57.8, 73.3)	66.7	(56.8, 75.4)	0.9088	0.8435
Traditional lottery	64.2	(56.0, 71.7)	61.3	(51.6, 70.2)	0.6448	0.7590
Instant games	42.0	(33.8, 50.7)	45.6	(36.3, 55.2)	0.5838	0.5160
Daily games	16.1	(10.6, 23.7)	20.4	(13.5, 29.6)	0.4175	0.0054
Raffles	48.9	(40.4, 57.5)	43.1	(34.0, 52.7)	0.3761	0.4676
Casino out of state	31.4	(24.0, 40.0)	33.4	(24.4, 43.8)	0.7636	0.8194
Casino in MA	NA		37.7	(29.2, 47.0)		
All Casino	31.4	(24.0, 40.0)	46.2	(36.8, 55.9)	0.0236	<0.0001
Sports betting	21.8	(15.7, 29.4)	19.2	(11.5, 30.1)	0.6540	0.7218
Private betting	18.8	(12.7, 27.1)	13.3	(7.3, 23.1)	0.3096	0.3414
Horse racing	5.3	(3.0, 9.2)	---		0.4050	0.8141
Bingo	***2.4***	***(1.0, 5.4)***	***5.4***	***(1.6, 16.3)***	0.3691	0.3225
Online	---		***2.2***	***(0.9, 5.3)***	0.8007	0.3389

RED indicates estimates are unreliable, relative standard error >30%
If cell size is 5 or less, results are set to dash (---)
CS indicates Chi-square test used
CMH indicates Cochran-Mantel-Haesznal test used

Table 9. Gambling Intensity: Targeted Springfield Survey (Weighted)

	Baseline (2015)			Follow-up (2019)			
	n		95% CI	n		95% CI	p-value[MH]
Total gambling expenditures:mean	1,127	-5547	(-12,257.4, 1,163.8)	1,129	-1169	(-2,400.2, 61.6)	0.8401
Total gambling expenditures:median	.	-14.0	(-43.4, 15.3)	.	-1.0	(-18.8, 16.9)	
Total IN and OUT of state casino expenditures:mean	1,064	-747.6	(-1,834.0, 338.9)	1,040	-1005	(-2,294.6, 284.9)	<0.0001
Total IN and OUT of state casino expenditures:median	.	-0.6	(-26.8, 25.6)	.	-0.7	(-64.1, 62.7)	
Total NON casino gambling expenditures:mean	1,062	-2619	(-6,745.5, 1,508.3)	1,039	-219.8	(-507.1, 67.6)	0.0288
Total NON casino gambling expenditures:median	.	-11.7	(-27.6, 4.2)	.	-10.1	(-22.1, 1.8)	
Max. freq. of gambling:mean	1,127	30.7	(25.2, 36.2)	1,132	34.0	(27.3, 40.8)	0.98
Max. freq. of gambling:median	.	4.0	(3.7, 4.4)	.	3.8	(3.5, 4.1)	
Number of gambling formats:mean	1,131	2.0	(1.8, 2.1)	1,134	2.2	(2.0, 2.3)	0.0320
Number of gambling formats:median	.	1.2	(1.0, 1.5)	.	1.4	(1.2, 1.7)	

Weighted N is the total number of respondents who answered the question weighted to the MA population
RED indicates estimates are unreliable, relative standard error >30%
If cell size is 5 or less, results are set to dash (---)
MH indicates that Mann Whitney U/ Wilcoxon rank sum test used
Expenditures are based on the entire population. Those who did not gamble had expenditures set to zero.

Table 10. PPGM (Collapsed): Targeted Springfield Survey (Weighted)

		Baseline (2015)			Follow-up (2019)			
		N	%	95% CI	N	%	95% CI	p-value[CS]
PPGM	non gambler	131,775	29.1	(25.4, 33.2)	116,705	25.6	(22.2, 29.4)	0.1993
	recreational gambler	266,215	58.8	(54.6, 63.0)	290,349	63.8	(59.6, 67.8)	0.1011
	at risk gambler, problem or pathological gambler	54,415	12.0	(9.3, 15.5)	48,204	10.6	(8.0, 13.9)	0.5114

Weighted N is the total number of respondents who answered the question weighted to the MA population
If cell size is 5 or less, results are set to dash (---)
CS indicates Chi-square test used

Table 11. PPGM: Targeted Springfield Survey (Weighted) by Gender

		Males					Females					
		BTPS_SP (2015)		FTPS_SP (2019)			BTPS_SP (2015)		FTPS_SP (2019)			
		%	95% CI		95% CI	p-value[CS]	%	95% CI	%		p-value[CS]	p-value[CMH]
PPGM 3 categories	non gambler	25.1	(19.1, 32.2)	22.4	(17.3, 28.6)	0.5770	32.2	(27.8, 37.1)	27.3	(23.1, 32.0)	0.1472	0.1715
	recreational gambler	58.1	(50.8, 65.1)	60.9	(53.8, 67.6)	0.5150	59.0	(54.1, 63.7)	65.4	(60.6, 69.9)	0.0505	0.0869
	at-risk/problem/pathological gambler	16.7	(11.7, 23.3)	15.7	(10.8, 22.4)	0.8362	8.0	(5.6, 11.3)	6.0	(4.3, 8.3)	0.2725	0.5052

CS indicates Chi-square test used
CMH indicates Cochran-Mantel-Haesznal test used

Table 12. PPGM: Targeted Springfield Survey (Weighted) by Age

		18-34					35-49				
		BTPS_SP (2015)		FTPS_SP (2019)		35-49	BTPS_SP (2015)		FTPS_SP (2019)		
		%	95% CI		95% CI	p-value[CS]	%	95% CI	%		p-value[CS]
PPGM 3 categories	non gambler	31.9	(23.7, 41.4)	36.5	(28.8, 45.1)	0.4525	31.5	(22.4, 42.3)	19.4	(13.4, 27.1)	0.0584
	recreational gambler	49.8	(40.5, 59.2)	51.1	(42.5, 59.7)	0.8467	62.0	(51.6, 71.4)	70.5	(61.3, 78.2)	0.2212
	at-risk/problem/pathological gambler	18.3	(11.5, 27.7)	12.4	(6.7, 21.6)	0.2835	*5.7*	*(2.9, 10.9)*	*9.6*	*(5.1, 17.3)*	0.2846

		50-64					65+					
		BTPS_SP (2015)		FTPS_SP (2019)			BTPS_SP (2015)		FTPS_SP (2019)			
		%	95% CI		95% CI	p-value[CS]	%	95% CI	%		p-value[CS]	p-value[CMH]
PPGM 3 categories	non gambler	21.8	(16.5, 28.1)	18.6	(13.0, 26.0)	0.4636	29.1	(23.8, 35.1)	23.0	(17.2, 30.1)	0.2253	0.1644
	recreational gambler	62.5	(55.3, 69.2)	70.1	(61.6, 77.3)	0.1752	62.3	(55.6, 68.5)	64.1	(56.2, 71.3)	0.4649	0.0927
	at-risk/problem/pathological gambler	15.2	(10.2, 21.9)	11.2	(6.5, 18.6)	0.3380	*8.2*	*(4.4, 14.7)*	9.6	(5.9, 15.3)	0.6229	0.5765

RED indicates estimates are unreliable, relative standard error >30%
If cell size is 5 or less, results are set to dash (---)
CS indicates Chi-square test used
CMH indicates Cochran-Mantel-Haesznal test used

Table 13. PPGM: Targeted Springfield Survey (Weighted) by Race/Ethnicity

		Hispanic/Black/Asian					White/Other					
		BTPS_SP (2015)		FTPS_SP (2019)			BTPS_SP (2015)		FTPS_SP (2019)			
		%	95% CI		95% CI	p-value[CS]	%	95% CI	%		p-value[CS]	p-value[CMH]
PPGM 3 categories	non gambler	40.2	(30.4, 50.9)	39.0	(31.9, 46.6)	0.8952	25.4	(21.6, 29.6)	18.6	(15.1, 22.8)	0.0194	0.1070
	recreational gambler	45.8	(35.9, 56.0)	44.7	(37.7, 51.9)	0.9117	62.4	(57.8, 66.9)	71.4	(66.3, 76.1)	0.0062	0.0509
	at-risk/problem/pathological gambler	13.0	(7.2, 22.3)	14.4	(9.7, 20.9)	0.7407	12.0	(8.9, 16.0)	9.3	(6.1, 13.7)	0.3054	0.5022

CS indicates Chi-square test used

CMH indicates Cochran-Mantel-Haesznal test used

NOTE: The same race/ethnicity groups used in other SEIGMA reports are used in this report to maintain consistency. Given the small proportion of individuals who identified as "Other" (8 out of 876 or 0.09% in the BTPS and 15 out of 693 or 2.16% in the FTPS), grouping these individuals with individuals who identified as White primarily highlights differences between Blacks, Hispanics and Asians, on the one hand, and Whites, on the other.

Table 14. PPGM: Targeted Springfield Survey (Weighted) by Education

		High School					College				
		BTPS_SP (2015)		FTPS_SP (2019)			BTPS_SP (2015)		FTPS_SP (2019)		
		%	95% CI		95% CI	p-value^{CS}	%	95% CI	%		p-value^{CS}
PPGM 3 categories	non gambler	33.5	(26.2, 41.6)	26.0	(19.6, 33.6)	0.1832	26.6	(22.3, 31.5)	23.2	(19.4, 27.5)	0.2853
	recreational gambler	51.6	(43.6, 59.5)	56.8	(48.4, 64.9)	0.3349	61.1	(55.8, 66.1)	68.5	(63.9, 72.8)	0.0269
	at-risk/problem/pathological gambler	14.0	(8.9, 21.2)	15.3	(9.8, 23.0)	0.7451	12.3	(8.9, 16.7)	7.9	(5.7, 10.8)	0.0640

		Graduate School					
		BTPS_SP (2015)		FTPS_SP (2019)			
		%	95% CI	%	95% CI	p-value^{CS}	p-value^{CMH}
PPGM 3 categories	non gambler	23.3	(18.3, 29.1)	26.9	(20.3, 34.7)	0.4371	0.1625
	recreational gambler	71.7	(65.5, 77.2)	66.3	(58.4, 73.4)	0.2688	0.0937
	at-risk/problem/pathological gambler	5.0	(2.9, 8.5)	6.8	(3.8, 11.9)	0.4515	0.5791

CS indicates Chi-square test used

CMH indicates Cochran-Mantel-Haesznal test used

Table 15. PPGM: Targeted Springfield Survey (Weighted) by Income

		<$50K					$50-99K				
		BTPS_SP (2015)		FTPS_SP (2019)			BTPS_SP (2015)		FTPS_SP (2019)		
		%	95% CI		95% CI	p-value[CS]	%	95% CI	%		p-value[CS]
PPGM 3 categories	non gambler	38.0	(31.5, 44.9)	25.5	(20.0, 31.9)	0.0076	22.1	(15.9, 29.8)	23.1	(17.2, 30.2)	0.8229
	recreational gambler	46.4	(39.8, 53.1)	59.1	(51.8, 66.0)	0.0117	65.4	(57.1, 72.9)	68.0	(60.2, 75.0)	0.5928
	at-risk/problem/pathological gambler	14.8	(10.2, 21.0)	14.4	(9.4, 21.6)	0.9390	12.5	(7.5, 20.0)	*8.4*	*(4.6, 14.9)*	0.3141

		$100K +					
		BTPS_SP (2015)		FTPS_SP (2019)			
		%	95% CI	%	95% CI	p-value[CS]	p-value[CMH]
PPGM 3 categories	non gambler	14.8	(10.5, 20.4)	17.0	(11.7, 24.1)	0.5802	0.2385
	recreational gambler	73.2	(65.1, 79.9)	72.0	(62.9, 79.7)	0.8447	0.1925
	at-risk/problem/pathological gambler	12.1	(7.0, 20.1)	*11.0*	*(5.7, 20.0)*	0.8174	0.7406

RED indicates estimates are unreliable, relative standard error >30%
CS indicates Chi-square test used
CMH indicates Cochran-Mantel-Haesznal test used

Table 16. Awareness of Media Campaigns and Programs: Targeted Springfield Survey (Weighted)

	Baseline (2015)			Follow-up (2019)			
	N	%	95% CI	N	%	95% CI	p-value
Aware of Media campaigns to prevent problem gambling	214,757	47.9	(43.7, 52.0)	146,906	32.1	(28.5, 36.0)	<0.0001 CS
Aware of Programs to prevent problem gambling	86,835	19.2	(16.2, 22.6)	66,762	14.7	(12.2, 17.6)	0.0340 CS

Weighted N is the total number of respondents who answered the question weighted to the MA population
If cell size is 5 or less, results are set to dash (---)
CS indicates Chi-square test used

Table 17. Treatment: Targeted Springfield Survey (Weighted)

		Baseline (2015)			Follow-up (2019)		
		N	%	% 95% CI	N	%	95% CI
Have you wanted help for gambling problems in the past year	No	10,642	81.0	(40.4, 96.4)	8,757	97.3	(81.8, 99.7)
	Yes	***2,492***	***19.0***	***(3.6, 59.6)***	***242***	***2.7***	***(0.3, 18.2)***
Have you sought help for gambling problems in the past year	No	2,269	91.0	(41.0, 99.3)	242	100.0	(. , .)
	Yes	***223***	***9.0***	***(0.7, 59.0)***	**0**		

Weighted N is the total number of respondents who answered the question weighted to the MA population
Note: RED indicates estimates are unreliable, relative standard error >30%
If cell size is 5 or less, results are set to dash (---)