UNSAL, NILAY, Ph.D. Three Essays on Evaluating Effectiveness of Workplace Wellness Programs. (2018)
Directed by Dr. Jeremy W. Bray. 195 pp.

Workplace wellness programs (WWPs) target enhancing employees' physical and mental well-being and provide potential health and economic benefits to workplaces. Workplaces have increasingly been adopting WWPs due to these perceived benefits. National and international policies have encouraged the use of WWPs to promote public health. The broad awareness and applications of WWPs require us to ask if and how we should evaluate their performance. Despite the pervasiveness, there have been disagreements on health improvements and potential financial savings of WWPs. The debate results from the difficulty to analyze these programs. Program design, program participation, measurement strategies, and statistical analyses are the main fundamental challenges that cause the difficulties. Many studies have discussed these challenges but have not proposed sufficient rigorous evaluation to validate the findings. This research advances the literature by characterizing the inherent methodological challenges in WWP evaluations, by finding solutions to some of these challenges, and by validating WWPs effectiveness with more comprehensive approaches. This dissertation fill the need for a rigorous evaluation by (1) evaluating the relationship between the methodological quality of studies and findings on returns of WWPs, (2) exploring reasons for adoption of WWPs and analyzing the assumption that financial savings drive the adoption of WWPs, and (3) estimating return on investment (ROI) of a WWP for a small nonprofit organization as well as evaluating the WWP outcomes with a more comprehensive approach.

THREE ESSAYS ON EVALUATING EFFECTIVENESS OF WORKPLACE WELLNESS PROGRAMS

by

Nilay Unsal

A Dissertation Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Greensboro
2018

Approved by

_____
Committee Chair

To Gokhan Karaman

APPROVAL PAGE

This dissertation written by Nilay Unsal has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair _____
Jeremy W. Bray

Committee Members _____
Albert Link

_____
Martijn van Hasselt

_____
Daniel L. Bibeau

_____
Date of Acceptance by Committee

_____
Date of Final Oral Examination

ACKNOWLEDGMENTS

I would like to acknowledge the great support of my dissertation advisor, Dr. Jeremy Bray, who has offered me much appreciated constructive suggestions, patient guidance, and enthusiastic encouragements. I also would like to express my gratitude to my committee members, Dr. Albert Link, Dr. Daniel Bibeau, and Dr. Martijn van Hasselt, for their constructive critiques, thoughtful comments, and encouragements.

In addition, I would like to acknowledge the contributions of past and present graduate students at UNC–Greensboro with particular thanks to Kelsi Hobbs, Dr. Gracielee Weaver, Dr. Justin Larson, Lorissa Pagan, Seth Tousey, Ian Mikkelsen, and Samar Alwehaibi, who all supported me with their friendships, suggestions, and encouragements. Thank you as well to Michael Pittard, Kelsi Hobbs, and Gracielee Weaver who helped edit this dissertation.

I would like to extend my sincere gratitude to the Well-Spring Retirement Community, especially to Garrett Saake, for providing me access to their data, which was vital to my dissertation. My thanks are also extended to Dr. Stephen Holland, Dr. Christopher Swann, and Dr. Martin Anderson for their help in offering me resources in completing my dissertation.

Lastly, I would like to thank my husband, Gokhan, my parents Mehmet Ali and Serife, and my sisters Dilek and Sevval, for their ongoing patience and inspiration through this adventure.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER I

INTRODUCTION

Chronic diseases have become the leading cause of disability and death, accounting for 60 percent of the deaths worldwide (Bloom et al., 2011). Global economic burden of major non-communicable chronic diseases was estimated $5.8 trillion for 2010 and projected $8.2 trillion for 2030 by the World Economic Forum. A report from Milken Institute showed that the total cost of chronic diseases in the U.S. health care system was $3.7 trillion in 2016 (Waters and Graf, 2018). The high prevalence and costs of chronic illnesses, such as heart disease, stroke, high blood pressure, diabetes, and obesity, show the importance of preventing diseases and promoting health, as well as treating diseases (WHO, 2005; Chénier. et al. 2012; Guazzi et al., 2013; Mattke et al., 2013; Edington et al., 2014; Benjamin, 2016; National Center for Health Statistics, 2016).

Workplace wellness programs (WWPs), which are employer-sponsored programs to promote health-related behaviors of employees, have become common intervention tools to address the rising prevalence and costs of chronic conditions, and advocate public health (Mattke et al., 2013). WWPs have widely recognized to prevent the spread of chronic diseases not only in Western societies, such as United States and Western Europe, but also in many other countries which are WHO members (Burton, 2010). Worksites are ideal places to reach out to a large portion of the population for health interventions (Cohen, 1985; Anderko et al.; 2012; Rongen et al., 2013). Work- related

factors, such as work environment and social relationships, impact individuals' ability to adopt healthy lifestyles (WHO, 2002).

Interventions such as WWPs is not a new notion. Public and private sectors have used workplace interventions as policy tools to improve employee health and productivity for decades (De Greef and Van den Broek, 2004; Spence, 2015). However, the concepts of health and wellness have evolved over time. The earliest interventions focused on worker productivity changes and improved working conditions, such as lighting, working hours, and rest breaks (Hawthorne effect). These efforts go back to the 1920s and 1930s (Jones, 1992). While those interventions may have had an indirect influence on the health of employees, the first intentional health-related programs to be implementing in workplaces was Employee assistance programs (EAPs). EAPs started as occupational alcoholism interventions that impacted worker productivity and can be traced back to the early 1940s in Western industrialized countries (Walsh, 1982).

The increasing prevalence and costs of chronic diseases initiated the health promotion and prevention (HPP) programs for employees starting in the 1970s (Cohen, 1985). Targeting individual risk factors using HPP programs was a response to increased health care costs due to chronic illnesses in the U.S. and increased illness- related productivity losses in Western Europe (Cordia et al., 2000; De Greef and Van den Broek, 2004). In the 1980s, employee wellness programs focused more on changing health behaviors and attitudes towards nutrition, weight management, alcohol and tobacco use, physical activity, and stress management (Cordia et al., 2000; Khoury, 2014).

The wellness perception of policy-makers has evolved from passive protection at early stages of occupational safety to now being recognized as an essential element of health development (Declaration, 1997). The wellness concept has started incorporating social and environmental components as well as behavioral changes since the late 1990s. Enhancing overall employee health beyond conventional occupational health and safety notions has become important for public health concerns. Today, WWPs are combined with occupational health and safety to enhance public health (WHO, 1997; Cordia et al., 2000; Warr, 2012).

WWPs have been popular and supported by international organizations and governments, especially starting in the late 1990s. The European Network for Workplace Health Promotion (ENWHP) and the World Health Organization (WHO) are the two most active institutions that support WWPs at the international level. The ENWHP is a network that assists organizations' corporate strategy to improve employee health and reduce the impact of work-related health issues (Guazzi et al., 2013). The World Health Organization (WHO) considers WWPs as the joint efforts of stakeholders and involves governments, nongovernmental organizations, and the private sector to promote health using workplace wellness (Declaration, 1997). In 2007, the WHO supported the Global Plan of Action (GPA) on workers' health, which examined the issue from a public health perspective, for the period of 2008-2017 (Burton, 2010).

The Affordable Care Act of 2010 (ACA), which is the largest national commitment to invest in wellness, is the U.S. government's strategy to use WWPs as a part of overall national health care policy (Anderko et al., 2012; Mattke et al., 2013).The

ACA provided technical assistance in terms of evaluating the effectiveness of WWPs and grants to small businesses in promoting health and preventing diseases. Prior to the ACA, the National Institute for Occupational Health and Safety (NIOSH) introduced the Total Worker Health (THW) program in 2003 as a commitment to protect and promote workers' health in the U.S. (Schill and Chosewood, 2013). The program established six broad priorities for the future research for WWPs: assessing intervention efficacy; focusing on population, job, and worksite characteristics; using proper study design (going beyond the reliance on randomized clinical trials); applying proper measures and metrics; studying sustainability and knowledge; and addressing global concerns (NIOSH, 2012).

Both the popularity and the amount of funds invested in these programs require researchers to establish these program's benefits to justify the use of WWPs as robust policy tools in public health. Evidence from the literature suggests that WWPs improve employee health by reducing modifiable risk factors, such as physical inactivity, tobacco use, unhealthy eating habits, obesity, high blood pressure, high blood glucose, and high cholesterol that could all cause chronic diseases (Meenan et al., 2010; Nyman et al., 2012; Goetzel et al., 2014; Dement et al., 2015; Newman et al., 2015). WWPs impact employee performance and turnover as well as employee health. Many studies have suggested that health improvements through WWPs increase health- related productivity by reducing absenteeism (absence from work due to sickness) and presenteeism (present at work but reduced productivity due to sickness) (Baker et al., 2008; Goetzel and Osminowzki, 2008; Trogdon, Reyes, and Dietz, 2009; Meenan et al., 2010; Yen et al.,

2010; Nyman et al., 2012; Goetzel et al., 2014). Previous research has mostly focused on the return on investment (ROI) of WWPs and suggested that companies that invest in WWPs could take advantage of financial savings (Cohen, 1985; Baicker, Cutler, and Song, 2010). Improvements in employee well-being and performance could decrease the organizational costs associated with health care utilization, high turnover, and health-related productivity losses (Goetzel and Ozminowski, 2008; Baicker et al., 2010; Henke et al., 2011; Goetzel et al., 2012; Nyman et al., 2012; Goetzel et al., 2014; Musich et al., 2014; Dement et al., 2015; Musich et al., 2015; Newman et al., 2015).

Purpose of the Study

The economic evaluation literature of WWPs still suffers from a lack of rigorous findings of program impacts (Hunnicutt and Leffelman, 2007; Lewis et al. 2014). Most of the prior studies have indicated the same limitations over the past three decades yet have not offered any effective methods to solve the issues. This lack of solutions results in little reliable evidence on effectiveness of WWPs that support employee-wellness related policies from the perspectives of employers and social policy-makers (Horwitz et al., 2013; Barbosa et al., 2015). The goal of this dissertation is to offer novel perspectives on how to improve the successful adoption and assessment of wellness programs in workplaces. The findings of this research assist both employers deciding on provisions of WWPs, as well as policy-makers supporting workplace wellness as a tool to promote public health.

This research advances the literature by characterizing the inherent challenges in WWP evaluations, by finding solutions to some of these challenges, and by validating

WWPs effectiveness with more comprehensive approaches. This dissertation fills the need for rigorous evaluations by providing three essays. The first essay, presented in Chapter II, is to understand the current state of workplace wellness. The findings establish the need to extend evaluation content beyond positive ROI and large for-profit companies. The second essay, presented in Chapter III, analyzes the assumption that financial savings drive the adoption of WWPs. This goal requires documenting to what extent ROI explains WWP adoption. The third essay, presented in Chapter IV, is designed to address issues raised in Chapter II by providing an evaluation of a WWP in a small nonprofit organization, chosen because where the literature fails to support the importance of WWPs is in organizations other than large and for-profit companies. Chapter IV addresses issues raised in Chapter III by examining the choice of WWP type along with the organization's goals and adoption reason for WWP.

Significance

This research provides several significant contributions to the field. First, Chapter II contributes to the field by identifying the common issues in WWP evaluations. This chapter offers solutions to the issues that can be fixed, such as needs for rigorous evaluations on small companies, statistical inference information, better reporting quality of studies, and independent evaluations.

Second, Chapter III contributes to the field by identifying the missing alignment in organizational values with WWP choices and WWP evaluation metrics. This is the first study that uses economic theory to model firm behavior when implementing and evaluating WWPs. An employer should choose a WWP where the chosen program has

the lowest opportunity cost in terms of achieving the company objectives. To measure the success given the cost, programs need to be evaluated based on the objectives for implementing. If an employer decides on implementing WWPs to receive positive returns, ROI should be the evaluation metric. If an employer decides on implementing WWPs to reduce number of injuries, ROI should not be the preferred metric. Currently, there is no theoretical or empirical studies on organizations' decision-making processes when implementing WWPs and WWP evaluation method that match with the organizations' objectives.

Third, Chapter IV contributes to the field by providing a rigorous example on a WWP evaluation. Chapter IV sets an example on aligning the company objectives for WWP implementation and evaluated outcomes as well as improve reporting quality for reliable evidence. WWPs are proposed as a strategy to improve employee well-being. However, there is not enough compelling evidence from observational studies. This chapter fills this gap by providing methods for evaluating a WWP for a small and nonprofit organization that potentially advances the use of the programs beyond large and for-profit organizations. Fourth, conflicts of interest in current WWP evaluations is a major issue. This study fills the need for independent research in the field and examines the investment decisions in WWPs from an objective and impartial perspective.

The remainder of this dissertation is organized as follows: Chapter II systematically reviews the economic evaluations of WWPs. This chapter examines the methodological quality of existing studies on the ROI of workplace wellness programs to identify limitations of existing studies. Chapter III examines whether positive ROI can

explain the reason for implementing WWPs. This chapter analyzes employers' reasons for WWP adoption to offer proper WWP components and evaluation methods. Chapter IV provides an evaluation for ROI of a wellness program in a long-term care company. This chapter presents a strategy for estimating the ROI of a workplace wellness program using a rigorous method. Chapter V concludes the dissertation with key findings, limitations, and future research avenues.

References

1. Aldana, S. G. (2001). Financial impact of health promotion programs: A comprehensive review of the literature. *American Journal of Health Promotion,* 15(5): 296–320.

2. Anderko L, Roffenbender JS, Goetzel RZ, Millard F, Wildenhaus K, et al. (2012). Promoting Prevention through the Affordable Care Act: Workplace Wellness. Prev Chronic Dis, 9.

3. Arrow, K. J. (2000). To profit or not to profit: The commercial transformation of the nonprofit sector. Cambridge University Press.

4. Baicker, K., Cutler, D., & Song, Z. (2010). Workplace wellness programs can generate savings. *Health affairs,* 29(2), 304-311.

5. Baker, K. M., Goetzel, R. Z., Pei, X., Weiss, A. J., Bowen, J., Tabrizi, M. J. & Thompson, E. (2008). Using a return-on-investment estimation model to evaluate outcomes from an obesity management worksite health promotion program. *Journal of Occupational and Environmental Medicine,* 50(9), 981-990.

6. Barbosa, C., Bray, J.W., Dowd, W.N., Mills, M.J., Kelly, E.L., Moen, P., & Wipfli, B. (2015). Return on investment of a Work-Family Intervention: Evidence from the Work, Family, and Health Network. *Journal of Occupational and Environmental Medicine,* 57(9), 943–951.

7. Benjamin, G. C. (2006). Putting the public in public health: New approaches. *Health Affairs*, *25*(4), 1040-1043.

8. Bertera, R. L. (1990). The effects of workplace health promotion on absenteeism and employment costs in a large industrial population. *American Journal of Public Health*, 80(9), 1101-1105.

9. Burton, J., & World Health Organization. (2010). WHO Healthy workplace framework and model: Background and supporting literature and practices.

10. Caloyeras, J. P., Liu, H., Exum, E., Broderick, M., & Mattke, S. (2014). Managing manifest diseases, but not health risks, saved PepsiCo money over seven years. *Health Affairs*, *33*(1), 124-131.

11. Chénier, L., Hoganson, C., & Thorpe, K. (2012). Making the business case for investments in workplace health and wellness. Conference Board of Canada.

12. Cohen, W. S. (1985). Health promotion in the workplace: A prescription for good health. *American Psychologist*, *40*(2), 213-216.

13. Cordia Chu, Gregor Breucker, Neil Harris, Andrea Stitzel, Xingfa Gan, Xueqi Gu, Sophie Dwyer. (2000). Health-promoting workplaces—international settings development. *Health Promot Int*, 15(2): 155-167.

14. De Greef, M., & Van den Broek, K. (2004). Making the case for workplace health promotion. *Analysis of the Effects of WHP*.

15. Declaration, J. (1997). Jakarta Declaration on Leading Health Promotion into the 21st Century. *Fourth International Conference on Health Promotion, Jakarta*.

16. Dement, J. M., Epling, C., Joyner, J., & Cavanaugh, K. (2015). Impacts of workplace health promotion and wellness programs on health care utilization and costs: results from an academic workplace. *Journal of Occupational and Environmental Medicine*, 57(11), 1159-1169.

17. Edington, D. W., Pitts, J. S., & Schultz, A. B. (2014). Health as an economic strategy: special emphasis on the workplace. *American Journal of Lifestyle Medicine*, *8*(5), 330-335.

18. European Network Workplace Health Promotion (ENWHP). (1997). *Luxembourg declaration on workplace health promotion in the European Union*. Bundesanstalt für Arbeitschutz und Arbeitsmedizin.

19. Gerard, K. (1992). Cost-utility in practice: a policy maker's guide to the state of the art. *Health Policy*, *21*(3), 249-279.

20. Goetzel, R. Z., & Ozminkowski, R. J. (2008). The health and cost benefits of work site health-promotion programs. *Annu. Rev. Public Health*, 29, 303-323.

21. Goetzel, R. Z., Tabrizi, M., Henke, R. M., Benevent, R., Brockbank, C. V. S., Stinson, K. & Newman, L. S. (2014). Estimating the return on investment from a health risk management program offered to small Colorado-based employers. *Journal of Occupational and Environmental Medicine*, 56(5), 554-560.

22. Guazzi, M., Faggiano, P., Mureddu, G. F., Faden, G., Niebauer, J., & Temporelli, P. L. (2014). Worksite health and wellness in the European Union. *Progress in cardiovascular diseases*, *56*(5), 508-514.

23. Henke, R. M., Goetzel, R. Z., McHugh, J., & Isaac, F. (2011). Recent experience in health promotion at Johnson & Johnson: lower health spending, strong return on investment. *Health Affairs*, 30(3), 490-499.

24. Horwitz, J. R., Kelly, D. R., DiNardo, J.E. (2013). Wellness incentives in the workplace: Cost savings through cost shifting to unhealthy workers. *Health Affairs*, 32(3), 468-476.

25. Hunnicutt, D; Leffelman, B. (2007). WELCOA's Well Workplace Initiative. *WELCOA's Absolute Advantage Magazine*, 6(10), 2-3.

26. Jones, S. R. (1992). Was there a Hawthorne effect? *American Journal of Sociology*, *98*(3), 451-468.

27. Kent, K., Goetzel, R. Z., Roemer, E. C., Prasad, A., & Freundlich, N. (2016). Promoting healthy workplaces by building cultures of health and applying strategic communications. *Journal of occupational and environmental medicine*, *58*(2), 114-122.

28. Khoury, A. (2014). The Evolution of Worksite Wellness. Corporate Wellness Magazine. Accessed May 23, 2017 from http://www.corporatewellnessmagazine.com/worksite-wellness/the-evolution-of/.

29. Kickbusch, I., & Payne, L. (2003). Twenty-first century health promotion: the public health revolution meets the wellness revolution. *Health Promotion International*, 18(4): 275-278.

30. Lewis, A. and Khanna, V. S. (2013, January 16) Is It Time to Re-Examine Workplace Wellness 'Get Well Quick' Schemes? Accessed January 28, 2017, from http://thehealthcareblog.com/blog/2014/12/12/the-wellness-industrys-terrible-horrible-no-good-very-bad-we/

31. Lewis, A. and Khanna, V. S. (2014, December 12). The Wellness Industries Terrible, Horrible No Good Very Bad Week. Accessed May 28, 2017, from http://healthaffairs.org/blog/2013/01/16/is-it-time-to-re-examine-workplace-wellness-get-well-quick-schemes/

32. Lewis, A., Khanna, V., and Montrose, S. (2014, November 25). Workplace Wellness Produces No Savings. Accessed January 28, 2017, from http://healthaffairs.org/blog/2014/11/25/workplace-wellness-produces-no-savings

33. Mattke, S., Liu, H., Caloyeras, J., Huang, C. Y., Van Busum, K. R., Khodyakov, D., & Shier, V. (2013). Workplace wellness programs study: Final report. *Rand Health Quarterly*, 3(2).

34. Mazur, B., & Mazur-Małek, M. (2017). Evaluating the Effectiveness of a Worksite Wellness Program. *International Journal of Contemporary Management*, *2016*(Numer 15 (4)), 77-89.

35. McCoy, M. K., Stinson, M. K., Scott, M. K., Tenney, M. L., & Newman, L. S. (2014). Health promotion in small business: a systematic review of factors influencing adoption and effectiveness of worksite wellness programs. *Journal of Occupational and Environmental Medicine/American College of Occupational and Environmental Medicine*, *56*(6), 579.

36. Meenan, R. T., Vogt, T. M., Williams, A. E., Stevens, V. J., Albright, C. L., & Nigg, C. (2010). Economic evaluation of a worksite obesity prevention and intervention trial among hotel workers in Hawaii. *Journal of occupational and environmental medicine/American College of Occupational and Environmental Medicine*, 52(Suppl 1), S8.

37. Merrill, R. M., Hyatt, B., Aldana, S. G., & Kinnersley, D. (2011). Lowering employee health care costs through the healthy lifestyle incentive program. *Journal of Public Health Management and Practice*, 17(3), 225-232.

38. Michaud, T. L., Nyman, J. A., Jutkowitz, E., Su, D., Dowd, B., & Abraham, J. M. (2016). Effect of Workplace Weight Management on Health Care Expenditures and Quality of Life. *Journal of Occupational and Environmental Medicine*, *58*(11), 1073-1078.

39. Mossink, J. and Nelson, D. I. (2002). Understanding and performing economic assessments at the company level (No. WHO/SDE/OEH/01.9). Geneva: World Health Organization.

40. Musich, S., White, J., Hartley, S. K., Bhattarai, G. R., Hawkins, K., & Ozminkowski, R. J. (2014). A More Generalizable Method to Evaluate Changes in Health Care Costs with Changes in Health Risks among Employers of All Sizes. *Population Health Management*, 17(5), 297-305.

41. National Center for Health Statistics US. (2016). Health, United States, 2015: with special feature on racial and ethnic health disparities.

42. National Institute for Occupational Health and Safety (NIOSH) (2012). *Research Compendium: The NIOSH Total Worker HealthTM Program: Seminal Research Papers* 2012. Washington, DC: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, DHHS (NIOSH) Publication No. 2012-146:1-214. Accessed January 28, 2017, from https://www.cdc.gov/niosh/docs/2012-146/pdfs/2012-146.pdf.

43. Nyman, J. A., Abraham, J. M., Jeffery, M. M., & Barleen, N. A. (2012). The effectiveness of a health promotion program after 3 years: evidence from the University of Minnesota. *Medical Care*, 50(9), 772-778.

44. O'Donnell, M. P. (2014). What is the ROI of workplace health promotion? The answer just got simpler by making the question more complicated. *American Journal of Health Promotion,* 28(6), iv-v.

45. Patient Protection and Affordable Care Act, 42 U.S.C. § 10408, (2010).

46. Patterson, M., Warr, P., & West, M. (2004). Organizational climate and company productivity: The role of employee affect and employee level. *Journal of Occupational and Organizational Psychology, 77*(2), 193-216.

47. Reardon, J. (1998). The history and impact of worksite wellness. *Nursing Economics*, 16(3), 117.

48. Rongen, A., Robroek, S. J., van Lenthe, F. J., & Burdorf, A. (2013). Workplace health promotion: a meta-analysis of effectiveness. *American journal of preventive medicine, 44*(4), 406-415.

49. Sorensen, G., & Barbeau, E. (2004). Steps to a healthier US workforce: integrating occupational health and safety and worksite health promotion: state of the science. *National Institute of Occupational Safety and Health Steps to a Healthier US Workforce Symposium*.

50. Sorensen, G., & Barbeau, E. (2012). *The Research Compendium: The NIOSH Total Worker Health™ Program: Seminal Research Papers 2012* (pp. 2012-146). Research Report. Washington, DC: US Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health.

51. Spence, G. B. (2015). Workplace wellbeing programs: if you build it they may NOT come... because it's not what they really need! *International Journal of Wellbeing*, 5 (2), 109-124.

52. Schill, A. L., & Chosewood, L. C. (2013). The NIOSH Total Worker Health™ program: an overview.

53. Trogdon, J., Finkelstein, E. A., Reyes, M., & Dietz, W. H. (2009). A return-on-investment simulation model of workplace obesity interventions. *Journal of Occupational and Environmental Medicine,* 51(7), 751-758.

54. van Dongen, J. M., van Berkel, J., Boot, C. R., Bosmans, J. E., Proper, K. I., Bongers, P. M., ... & van Wier, M. F. (2016). Long-term cost-effectiveness and return-on-

investment of a mindfulness-based worksite intervention: results of a randomized controlled trial. J*ournal of Occupational and Environmental Medicine*, *58*(6), 550-560.

55. Yen, L., Schultz, A. B., Schaefer, C., Bloomberg, S., & Edington, D. W. (2010). Long-term return on investment of an employee health enhancement program at a Midwest utility company from 1999 to 2007. *International Journal of Workplace Health Management*, 3(2), 79-096.

56. Walsh, D. C. (1982). Employee assistance programs. *The Milbank Memorial Fund Quarterly. Health and Society*, 492-517.

57. Warner, K. E., Wickizer, T. M., Wolfe, R. A., Schildroth, J. E., & Samuelson, M. H. (1988). Economic implications of workplace health promotion programs: review of the literature. *Journal of Occupational and Environmental Medicine*, 30(2), 106-112.

58. Warner, K. E. (1987). Selling health promotion to corporate America: Uses and abuses of the economic argument. *Health Education & Behavior*, 14(1), 39-55.

59. World Health Organization (WHO) (1997). WHO's Global Healthy Work Approach. Division of Health Promotion, Education & Communication and Office of Occupational Health, Geneva.

60. World Health Organization, Public Health Agency of Canada, & Canada. Public Health Agency of Canada. (2005). Preventing chronic diseases: a vital investment. World Health Organization.

61. World Health Organization. (2005). Participants at the 6th Global Conference on Health Promotion. The Bangkok Charter for health promotion in a globalized world. Geneva, Switzerland: World Health Organization.

CHAPTER II

A SYSTEMATIC REVIEW OF ECONOMIC EVALAUTIONS OF WORKPLACE WELLNESS PROGRAMS

Introduction

An extensive empirical body of literature suggests that workplace wellness programs (WWPs) improve employee health and work performance (Goetzel and Ozminowzki, 2008; Trogdon, Reyes, and Dietz, 2009; Meenan et al., 2010; Henke et al., 2011; Goetzel et al., 2014; Newman et al., 2015). The economic evaluation literature of WWPs has mainly focused on the employer's perspective. Therefore, cost-savings and return on investment (ROI) from health and performance improvements have been the focus of much of this literature (Baker et al., 2008; Yen et al., 2010; Nyman et al., 2012; Musich et al., 2014; Dement et al., 2015). Based on this literature, many agencies, such as the US government and the European Network for Workplace Health Promotion, advocate for the widespread adoption of WWPs, particularly for the purpose of reducing health care costs and productivity losses.

Over the past decade, however, new criticism of this literature argues that the expected cost savings may not materialize (Lewis and Khanna, 2013, 2014). In particular, recent commentators assert that the literature suffers from a lack of rigorous evaluation. There is not enough reliable evidence on WWP effectiveness, in terms of delivering cost-savings or positive ROI, to support policies promoting the adoption of WWPs. In this

study, we present a systematic review of WWP ROI studies to both assess the evidence supporting WWP policy efforts and to identify the common methodological challenges in this literature.

Because a prior review suggested that studies with greater methodological rigor yield lower ROI estimates (Baxter et al., 2014), a primary goal of this systematic review is to determine if higher methodological quality is associated with lower ROI estimates for WWPs. To accomplish this goal, we used a broader quality index that contains measures relevant to common methodological limitations noted in criticism of the WWP economic evaluation literature. A lack of randomization is one such issue that decreases the study rigor. In general, the nature of WWP participation is nonrandom where employees voluntarily participate, such as when voluntary nature has been ensured by regulations such as the Affordable Care Act of 2010 (ACA), the Genetic Information Nondiscrimination Act of 2008 (GINA), the Health Insurance Portability and Accountability Act of 1996 (HIPAA), and the Americans with Disabilities Act of 1990 (ADA). Another common criticism is the lack of proper statistical methods to estimate costs, savings, and ROI. Lastly, potential conflict of interest is also a common criticism. Conflicts of interest could arise due to two main reasons: the WWP provider and the evaluator are the same or the evaluator has a financial interest in the WWP.

Beyond methodological rigor, we also explore other potential limitations of the literature that are less commonly noted. First, the literature does not use a common definition of ROI, leading to inaccurate information on returns being positive and limiting the validity of comparisons across studies. Second, statistical inference

information for ROI, such as confidence intervals and standard errors, is often absent such that formal meta-analyses cannot be performed. Third, large companies are overrepresented in the literature and other organizational characteristics such as country, industry, objectives for WWP adoption, and WWP provider are important elements that could impact the outcomes yet are reported inconsistently in the literature. Fourth, WWPs do not have a standard definition, which means that a disease management component is occasionally included, or prevention programs are labeled as wellness. Disease management and wellness target different outcomes, thus the content has different impact on ROI.

This systematic review provides insights into the main challenges inherent in the economic evaluation literature. Some of these issues cannot be solved due to legal or practical issues, such as providing more randomized studies and collecting certain individual data. However, research can offer analysis methods that could improve the study rigor that could better support meta-analyses. Furthermore, we confirm that large companies that adopt WWPs specifically in search of cost savings are heavily represented in the peer-reviewed literature of economic evaluations of WWPs. The literature needs to extend evaluations beyond positive ROI and large organizations to reflect more diverse employers' motivation for adopting WWPs correctly and have impact on decision-making process more effectively.

Methods

This systematic review was conducted using a predetermined protocol based on the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA)

guidelines. We identified peer reviewed articles in PubMed, EconLit, Proquest Central, and Scopus. The initial search was conducted on June 4, 2017 and the last search was conducted on December 7, 2017. The search resulted in 78 potential articles from which 33 met eligibility criteria. In addition, we included 11 articles that were pulled from reference lists in articles found in the search.

The search parameters are listed below. The target population was one of the following: workplace, employee, worksite, or worker. The target intervention was wellness, health, health promotion, health prevention, or wellbeing. The target evaluation was economic evaluation including cost benefit, cost effectiveness, cost analysis, economic evaluation, economic analysis, or economic assessments. The outcome was return on investment. Table 2.1 presents a detailed list of the search terms for each database. We did not limit the year of publication. However, we excluded publications that were not an independent study (e.g., a review, simulation, or meta-analysis) and not in English. We also did not limit the search to specific diseases to include workplace wellness, health promotion and disease management programs. We initially reviewed the articles based on the title and abstract. The next step was sifting through the articles to verify that the included articles were relevant for a full text review. After the full-text review, we scanned the reference lists of all identified publications, including those from systematic reviews, meta-analyses, and other reviews to identify relevant citations.

Eligibility Criteria and Study Selection

The inclusion criteria in the title and abstract review step were determined using the search categories presented in table 2.1. The objective was to use the least amount of

restrictions to get a broad range of articles that presented an economic assessment of WWPs. We excluded articles that did not target employees or workplaces. We excluded articles that did not evaluate health or wellness programs. We also excluded articles that were related only to work process and not related to health behavior, such as occupational safety and health (OSH), ergonomics, and employee assistance programs.

In the full-text sift, we imposed the same restrictions on target population, intervention, and outcome analysis as in the title and abstract sift. In addition, we excluded articles that evaluated government-sponsored WWPs to maintain the focus on employer-relevant information. Because this review focused on economic outcomes, we excluded studies that did not conduct an economic analysis. Only peer-reviewed articles were included to analyze the validity of recent critiques of WWP ROI studies (Lewis and Khanna, 2013, 2014; Lewis, Khanna, and Montrose, 2014).

Return on Investment

ROI is the primary economic outcome of interest for this systematic review. There are two ways ROI findings were extracted from the selected articles. The first is the ROI estimate as reported, regardless of how it was measured. Reported ROI measures include true ROI, expressed either as a ratio or a percentage and measured as the ratio of net benefit (the difference between benefits and program costs) to program cost, which has a threshold for positive ROI of zero (Phillips and Phillips, 2007); the benefit to cost ratio, which has the threshold for positive ROI of one; or net benefit with positive ROI as savings exceeding program costs. Second is recalculated ROI using net benefit to cost ratio with the threshold for positive ROI as zero to ensure comparability across studies. If

the study did not report ROI as its finding but reported program costs and benefits, then ROI was calculated using net benefit to cost ratio. The key for consistency in this review is to unify all ROI calculations across studies. Thus, the recalculated ROI measure is the outcome of interest.

Monetized values were not adjusted to real values or discounted to have standard valuation across studies. Discounting would require extracting annual flow information for costs and savings, which was not possible for all of the papers in this analysis. We did not adjust the real values across studies for two reasons. First, studies with program and evaluation duration overlaps adjusted both costs and benefits for real values. Second, studies with program duration shorter than evaluation duration adjusted benefits using the price index of the program year.

We used ordinary least square (OLS) regression to examine the conditional mean of ROI estimates across values of the quality index, organization size, program component (wellness or disease management), evaluated outcomes (costs of health care, absenteeism, and productivity), publication year, and study duration. We used estimated coefficients only to examine the mean effects and not to imply any causal relationship. Because few of the source papers included the standard error of the ROI estimate, the regression has not been adjusted for source study sampling variation and so should not be considered a true meta-regression.

Quality Index

A primary focus of this paper is the relationship between the methodological quality of studies and ROI estimates. To measure the methodological quality, we

extracted information from the included articles using a methodological rigor rubric that was adapted from the following checklists: Adams (1992), Gerard (1992), Sacristan (1993), and Downs and Black (1998). The methodological rigor rubric, presented in Appendix table 2.A1, includes domains for article characteristics, reporting, internal validity, external validity, and statistical power. We used eighteen equally weighted items (eight items from reporting, eight items from internal validity, one item from external validity, and one item from power) to calculate an overall quality index for each included article.

Article Characteristics Domain

Article characteristics include the following items: author(s), year that the study was published, journal in which the study was published, content of the evaluated program, country of the study, company or the industry that the study was evaluated, number of participants and nonparticipants, and size of the companies. In addition, we extracted the authors' academic department(s) or research center(s) and the study funding agency to identify possible conflicts of interest.

Information on WWP programmatic content was used to classify the program into two main categories: disease management and wellness program. A program is classified as disease management if it targeted diagnosable diseases (i.e. asthma or diabetes). A program is classified as wellness if it targeted health risks or behaviors (i.e. smoking, exercise, or nutrition). Based on this classification, some programs had the same components, such as weight loss and exercise, but they were not classified in the same category due to the program's target outcome.

The size of the organization was determined based on either the number of employees that were eligible to participate to the program or the number of participants and nonparticipants. Some studies included multiple worksite with various sizes. For those, the size of the company is considered as not available if the study did not provide any specific information.

Information from the authors' conflict of interest acknowledgements, the funding sources of the study, and the department that conducted the study were used to evaluate potential conflict.

Reporting Domain

The reporting domain includes: objectives of the study, intervention, study sample, type of economic analysis, main outcomes, program cost, main findings, and statistical inference for the main outcomes. Each item can get a score of one if it was reported before the results section. We modified the condition of "reporting before results" for study sample, program costs, and statistical inference information, where each of these components can score one if they were reported anywhere in the article and zero otherwise. This modification helps to ensure that articles are not scored poorly due to different publication practices across disciplines.

The study sample component has three equally weighted subcomponents: study population, sample selection criteria, and analysis sample. The study population component was scored one if information on eligible employees was provided and zero otherwise. The sample selection component was scored one if information on how eligible employees selected into intervention group was explained and zero otherwise.

The analysis sample item was scored one if information on final analytic sample was provided and zero otherwise.

Internal Validity Domain

The internal validity domain includes the following items: description of main outcome measures, study design, outcome evaluation method, cost measures and valuation, population that subjects were recruited from, period that subjects were recruited in, adjustment for different follow-up lengths or periods for groups, and attrition from the study sample.

Description of main outcome measures, such as health care, absenteeism, and productivity, is scored one if evaluated outcomes and their measurements are clearly described and zero otherwise. Study design is scored 1 if the study is randomized, 0.75 if the study is quasi-experimental, 0.5 if the study is an observational cohort study with a control group, 0.25 if the study is an observational case study with a control group, and 0 if the study is an observational study without a control group.

Study design, which captures randomized, quasi-experimental, and observational studies, is an internal validity component that differs based on group (intervention and control) assignments and data collection. We classified studies as randomized if the study clearly stated the design and provides information on how the groups are assigned. We classified the study as quasi-experimental only if the paper clearly stated the design is quasi-experimental. Finally, we classified observational studies based on the use of administrative data and whether it is a cohort or a case study with or without a control group.

Outcome evaluation method has two equally weighted subcomponents dealing with selection bias and using proper methods to estimate outcome variables. The first component is to identify if any method was used to increase the strength of causal inference if the study is not randomized. It is one if there is no evidence of possible selection bias or if there is an effort to reduce the selection bias such as using matching methods or intention-to-treat approach and zero otherwise. The second component is to identify if a proper method was used to estimate the program outcomes. It is one if the methods deal with issues such as clustering in randomized studies, skewed data, count data, binary data, etc. and zero otherwise.

Cost measures and valuation have three equally weighted subcomponents describing and measuring intervention costs, discounting, and price adjustment. Describing and measuring intervention costs is scored one if direct measures (units and unit prices) were used or if a cost regression was used to estimate the marginal cost of implementing the intervention and zero otherwise. Discounting is scored one if the net present values for program cost and monetized outcomes were presented. Discounting is also scored one if there is no discounting, but the reason was explained, such as the study period being less than a year. Discounting is scored zero otherwise. Price adjustment is scored one if both program costs and monetized outcomes were adjusted for inflation. Price adjustment is also one if there is no adjustment, but the reason was explained, or if it is clear in the text that no adjustment was necessary due to the length of the follow-up period. Price adjustment is scored zero otherwise.

The population that subjects were recruited from is scored one if the participants and nonparticipants were selected from the same population and zero otherwise. The period that subjects were recruited in is one if participants and nonparticipants were selected from the same time period and zero otherwise. Adjustment for different follow-up lengths is scored one if the follow-up was the same for all groups or if different follow-up lengths were properly accommodated in the analysis. Adjustment for different follow-up lengths is scored zero if the differences in follow-up across groups was ignored. Attrition from the study sample is scored one if it was mentioned, explained, and addressed in the analysis. Attrition from the study sample is also scored one if it was mentioned, but it was not handled because the loss was too small to affect the main findings. Attrition is scored zero otherwise.

External Validity Domain

There is one item in this domain to measure whether the study findings apply to the population. A case study needs to be externally valid for generalizability matters. However, WWPs are unique to their workplaces, which is one of the inherent issues with WWP evaluation literature that cannot generally be improved. They are not likely to be generalizable to other workplaces even if the program and study sample have similar characteristics. Therefore, we considered a study externally valid if the analyzed subjects represent the population from which they were recruited.

Power Domain

There is one item in this domain to measure whether the study has sufficient statistical power. This item is scored one if any information related to a formal power

analysis was found. Finding information related to statistical power in WWP evaluations is not common. Articles with power information provided either information on statistical power to ensure identification of the program effect or whether there is a decrease in power due to sample size.

Additional Analysis Variables

Organization Size

Resources available for a wellness program and the impact of a wellness program might differ based on the company size. Thus, ROI findings are likely to vary with the size. In this analysis, organization size is classified in two categories: small (500 or less employees) and large (more than 500 employees). Size is the only company characteristic used in the analysis due to the lack of other information across studies.

Program Component and Evaluated Outcomes

Publications were categorized into wellness and disease management programs that target different outcomes. The aims were to identify the composition of disease management and wellness programs among the selected articles and to analyze whether ROI results significantly differ based on these components.

Furthermore, studies were categorized based on the cost components included in the ROI analysis, such as costs of health care, absenteeism and productivity. Health care included pharmaceutical claims and medical claims of inpatient, outpatient, and emergency room visits. Absenteeism included work lost days, sickness absence days, disability days, or time away from work. Productivity mainly included presenteeism as

well as work performance. The aim is to analyze whether ROI results vary based on what was included in the benefits.

Publication Year and Study Follow-up Length

Publication year and the follow-up length are other two factors that might impact the ROI findings of studies. For the analysis, we used a year dummy variable that is one if the study was published in 2011 or after, and zero otherwise. The year 2011 was chosen to capture the potential impacts of the Affordable Care Act of 2010 on any publications' findings.

Study follow-up length matters to analyze short- and long-term effects of evaluated programs. We used a dummy variable that is one if the study follow-up was three years or longer, and zero otherwise. The three-year study duration was chosen based on the sample size.

Sensitivity Analyses Measures

Reduced Quality Index

The reduced quality index includes the reporting and internal validity domains of the rigor rubric. All sixteen items are equally weighted. There are two reasons why we excluded the external validity and power items. First, both measures have limited conceptual relevance for WWP ROI studies, as explained in the Data section. Second, the primary analyses suggest that the excluded domains have limited empirical relevance compared to the included domains. The aim of the sensitivity analysis is to analyze if the mean effects for ROI are sensitive to excluded domains.

Study Design

Although the overall quality index includes the study design item, study design is likely to have a large impact on ROI findings individually. Therefore, a separate analysis is needed to identify that impact. In randomized studies, the researcher randomizes intervention and control groups from either eligible employees or employees who responded to a program invitation. Although assignment to groups was random, nonrandom selection into the group of eligible employees or invitees was possible as a result of the nature of participation in WWPs. In quasi-experimental studies, the researcher assigns the groups using a nonrandom procedure. In observational studies, the researcher allows employees to self-select into program participation, and often uses administrative data related to a wellness program and its outcomes.

Estimation Methods for Monetized Outcomes (Benefits)

Outcome estimation is an internal validity item that determines whether appropriate statistical methods were used to estimate costs and benefits. The statistical techniques must be appropriate to the data, such as if the data were not normally distributed (skewed data) or if nonparametric approaches were needed. This item is likely to have a high impact on ROI findings.

Measuring Costs

Publications are scored based on how much detail they provided on program costs and whether benefits and costs were discounted and adjusted for inflation when the follow-up period is over a year. Although measuring program costs is a quality index item, a separate analysis was needed to evaluate whether more detailed information on

costs does indeed impact the ROI findings. This separate analysis only includes the item of whether the study provided detailed program cost information. Discounting or real value adjustments were not included because monetized values were not adjusted to real values or discounted in this analysis to have standard valuation across studies.

Additional Analysis

Conflict of Interest

Lack of independent studies in the economic evaluation literature of WWPs has been one of the most persistent critiques of the reliability of positive ROI findings. We therefore examine the distribution of articles that are independent without funding, independent with funding (where the funder was not the organization whose wellness program was evaluated), and not independent with funding (where the funder was the organization whose wellness program was evaluated). The aim is to show how many articles are independent and how many of them could potentially suffer from a conflict of interest, which might lead to biased findings for the evaluated program.

Results

Study Selection and Characteristics

For this systematic review, 466 unduplicated articles were identified and abstracted for further review, as shown in figure 1. Of these, 78 articles met the inclusion criteria for the title and abstract screening. Thirty-three of those articles were selected for full text review. In addition, 11 articles were included from the publications' reference lists, resulting in a total of 44 unique publications included in the review.

Tables 2.1-2.3 provide details for the key characteristics of included publications sorted by study design. Of all 44 publications, 13 were randomized studies, 4 were quasi-experimental studies, 3 were observational cohort studies with a control group, 16 were observational case studies with a control group, and 8 were observational studies without a control group. Twenty-eight studies evaluated only wellness programs, 5 evaluated only disease management programs, and 11 evaluated both wellness and disease management programs. Thirty-four studies were conducted in companies in the United States, 8 in the Netherlands, and 1 each in Japan and Germany. All studies conducted in the Netherlands and Germany were randomized studies. 23 articles were published in the *Journal of Occupational Environment and Medicine*, 3 were published in the *American Journal of Health Promotion*. Primary authors of 19 articles authored only one paper in the review. Fifteen primary authors have two or more papers in the review.

Half of the included articles were published after 2010 (2011 and after). The earliest publication was 1984 and the latest publication was 2017. The program start year differs from the publication year. The earliest program year goes back to 1977 and the latest program start year is 2013. Study follow-up length varies between 6 months to 15 years. Half of the included articles were published after 2010 and most of the articles have follow-up lengths of one, three, or five years.

Only 3 out of 44 publications evaluated a program in a small company. Of 44 publications, 14 provided clear information on which organizations' wellness programs were evaluated. Most of the studies stated the economic analyses were done from the

employer's perspective, however none of the studies reported the actual objective of organizations for adopting WWPs.

Of the 44 included articles, there is no information on funding source or conflicts of interest from 2 articles. Of the 42 articles with conflict of interest information, 24 had a statement of no authorial conflicts of interest. However, 10 out of 24 were funded by organizations, such as PepsiCo, Aetna, and University of Minnesota, whose wellness programs were being evaluated at the time of the study. In addition, 8 more articles did not acknowledge any conflicts of interest but had potential conflicts because they were funded by the same organization whose WWP was being evaluated. As a result, 18 (41%) included articles had potential conflicts of interest.

Analysis Sample for ROI Analysis

Twenty-eight studies reported an ROI outcome as shown in tables 2.4 and 2.5. Out of the 28, only 9 used net benefit to cost ratio as the ROI measure, as shown in tables 2.4 and 2.5. Of the 28, 3 of them used net benefit to cost ratio as percentage to report ROI in percent. More than half of the studies reported ROI as benefit to cost ratio, which leads to a higher ROI finding by construction. Two articles did not provide any information on the ROI formula they used. Moreover, only 4 out of 28 studies, which were randomized, reported confidence interval information for the ROI estimate. Two of them found positive ROI which were not significant.

We recalculated the ROI of 26 studies. We excluded 2 articles that did not have sufficient program cost information to perform the recalculations. In addition, we calculated the ROI of 3 studies, which did not provide ROI as outcome, using reported

program benefits and costs. Therefore, the final sample for the ROI regression analyses includes 29 articles with recalculated ROI outcomes.

Table 2.6 summarizes information on recalculated ROI, quality indices, and the rubric domains for the analysis sample. The mean of recalculated ROI is 0.68 with a minimum of -12.61 and a maximum of 10.17. The mean of quality index with all 4 domains is 11.97 points with a minimum of 5 and maximum of 18 points. The mean of the reduced quality index with the reporting and internal validity domains is 11.66 points, with a minimum of 5 and maximum of 16.

Table 2.7 summarizes the characteristics of the articles used in the regression analyses and some of the internal validity components that we will detail in the sensitivity analyses. Of 29 publications, only 3 of them were conducted in small companies. Twenty studies evaluated only wellness programs, 4 evaluated only disease management programs, and 5 evaluated both wellness and disease management programs. Seven studies included only health care costs savings (losses) in the ROI, 7 included only absenteeism cost savings (losses), 4 included both health care and absenteeism cost savings (losses), and 11 included any combination of cost savings (losses) that included productivity. There are 17 studies published after 2010 and 18 studies with a follow-up length of at least 3 years. Of 29 articles, 6 were observational studies without a control group (base group), 9 were observational studies with a control group, 4 were quasi-experimental studies, and 10 were randomized studies. Only 7 publications used proper statistical techniques to accommodate features of the study data. Nine publications

provided direct measures (units and unit prices) for program costs or cost regression to identify marginal costs.

Regression Analyses

We used OLS regressions to estimate the conditional mean of ROI estimates across studies. Coefficients from the OLS regressions should only be interpreted as differences in the conditional mean and not as causal effects. Although we present standard errors for regression coefficients, these standard errors do not account for the underlying sampling variation in the source studies because only 4 studies provided that information. Therefore, these regressions do not constitute a formal meta-analysis.

Table 2.8 shows the results of the main regression analyses. Although none of the results are statistically significant, columns 1 through 6 nonetheless provide useful information on the differences in ROI across individual study characteristics. Higher quality studies reported lower ROI. Small company evaluations produced lower ROI. Studies that evaluated disease management (base group) produced higher ROI compared to evaluations that contained a wellness component. Studies that included only cost of absenteeism or both costs of absenteeism and health care reported lower ROI compared to studies that included only cost of health care (base group). Studies that included any cost of productivity also reported lower ROI estimates compared to studies that included only the cost of health care. Studies that included productivity costs have the lowest ROI estimates. Studies published after 2010 reported lower ROI compared to studies published 2010 and before. Studies with a follow-up length of more than 3 years produced lower ROI compared to studies with follow-ups of 3 years or less. Column 7

controls for all components simultaneously. The sign of the coefficients and significance do not change except for the small company indicator.

The impacts of quality index, size indicator, and disease management and wellness program indicators are as expected. One would presume that adding more benefit components or lengthening the follow-up period could increase ROI due to potential increases in cost savings, yet the findings here suggest the opposite.

Sensitivity Analyses

Impact of Outliers

The highest ROI value is 10.17 (Noben et al., 2015) and the lowest ROI value is -12.61 (van Holland et al., 2017). Both studies are randomized and evaluated programs in large organizations. These ROI values are extreme outliers relative to the interquartile range of ROI estimates (see Mann (2007), p. 117 for the definition of extreme outlier used here). After removing these outliers from the recalculated ROI, the sign and relative magnitude of mean effects remain the same, except for study follow-up length as shown in Appendix 2.A2, columns 1 through 6. When all study characteristics are included in the regression, the signs of quality index, company size indicator (small), cost of only absenteeism, and publication year indicator changed from negative to positive. The results may be sensitive to the outliers in ROI, but none of these results are statistically significant (Appendix 2.A2).

Reduced Quality Index

Table 2.9 shows the ROI regression results using the reduce quality index. The findings, which are presented in table 2.9 columns 1 and 2, are similar to the findings

when the full quality index was used. In addition, table 2.9 columns 3 and 4 show the mean effects of the two highest impact domains (reporting and internal validity) separately. Studies with higher reporting scores reported lower ROI compared to studies with lower reporting scores. The effects are significant at 5% level. Studies with higher internal validity scores reported higher ROI compared to studies with lower internal validity scores. The effect on column 3 is significant at 10% level. The rest of the findings for other control variables are similar to the main analysis.

After removing outliers from the recalculated ROI, the sign and relative magnitude of mean effects remained the same, except for disease management and wellness program indicator, absenteeism cost, year indictor, and evaluation duration as shown in Appendix 2.A3 columns 3 and 4. In addition, the magnitudes for reporting variable decreased and lost significance. Mean effects are sensitive to extreme outliers in ROI (Appendix 2.A3).

Mean Effects on ROI Using Internal Validity Items of Study Design, Estimation Methods, and Program Cost Valuations

Table 2.10 shows the regression results for the internal validity items separately. Observational studies with a control group reported lower ROI compared to the observational studies without a control group. Similarly, randomized studies reported lower ROI compared to any other study design. ROI goes down if the design is more rigorous, except for the quasi-experimental studies. The classification of quasi-experimental design depends on whether the study clearly reported the design as quasi-experimental. Quasi-experimental studies might be misclassified due to missing

information on design where the design relied on regression models to compare outcomes for intervention groups (Musich et al., 2015) or different outcomes were analyzed with different designs (Grossmeier et al., 2013).

Studies that used proper estimation methods, as described in the Methods section, reported higher ROI compared to the studies that did not. Studies that provided detailed information on program costs reported higher ROI compared to studies that did not. Table 2.10 column 4 shows the mean effects when all these components are controlled. The direction of the coefficient and significance did not change, and the positive impact of detailed program costs had the highest impact on ROI estimates.

After removing the outliers, the sign and the relative magnitude of the coefficients remained the same except for clearly described program costs as shown in Appendix 2.A4. Studies that provided detailed information on program costs reported lower ROI findings compared to other studies. Program cost valuation is sensitive to outliers in ROI.

Additional Analyses

Conflicts of Interest

Of the 29 articles in the ROI regression analyses, one did not provide sufficient information to determine conflicts of interest. Of the remaining 28 articles, 14 had a statement of no authorial conflict of interest. However, 4 out of those 14 articles were funded by organizations whose wellness programs were evaluated. In addition, 5 more articles that did not acknowledge any potential conflict of interest had a potential conflict for the same reason.

Table 2.10 column 5 presents the mean effects of conflict of interest on ROI. Articles with potential conflict of interest reported higher ROI. The findings remain the same when the outliers in ROI were removed from the sample (Appendix 2.A4).

Discussion

This systematic review suggests that higher methodological quality results in lower ROI of workplace wellness programs, corroborating an earlier systematic review (Baxter et al., 2014). Studies with higher quality made use of methods that could reduce the positive bias in ROI. Study design item of internal validity domain potentially had the highest impact on ROI findings due to selection bias in program outcomes. Moreover, study design influences the methods for data collection and evaluation that could impact all the other items in the internal validity domain. Self-selection into the program due to voluntary participation might cause overvalued program benefits. Missing details in program costs, such as units, unit prices, and opportunity costs, might also cause undervalued program costs. Positive bias in benefit valuations and negative bias in cost valuations could lead to positive bias in ROI. Thus, we tested the findings based on quality index components in details to understand the relationship between study quality and ROI findings.

We used the reduced quality index that includes reporting and internal validity domains but excluded external validity and statistical power domains. The mean effects remained similar to what was found using quality index. We also used the reporting and internal validity domains separately instead of a quality index. The negative association between study quality and ROI was dominated by the reporting quality of the studies.

Higher internal validity scores were associated with higher ROI. Recent critiques of WWP ROI literature suggest that studies with greater internal validity should yield lower ROI estimates. To better understand the positive association of internal validity, we looked at internal validity items of study design, estimation methods, and program costs separately. Impacts of study design were as expected, where more rigorous design resulted in lower ROI. Studies that used appropriate outcome estimation methods, such as methods to deal with skewed or count data, reported higher ROI. The sign and magnitude of this effect might depend on valuation of benefits. Even if the methods control for non-normality; health care, absenteeism, and productivity costs could be greatly skewed due to high outliers. In addition, these costs could show great variability at the organization and employee levels in terms of industry, wages, fulltime status, and department.

Studies with more detailed program cost information that contained units, unit prices, and marginal cost estimations were expected to have lower ROI, due to higher estimates of program cost. However, the mean effects showed that studies with more detailed program cost reported higher ROI. The positive sign might be a result of high benefit levels in those studies. Third, we ran the analyses without extreme outliers. Although the results for quality indices were not sensitive, reporting and internal validity domains, as well as internal validity items, were sensitive to extreme outliers.

Results from the reporting and internal validity domains suggest two important conclusions. First, the negative impacts of quality indices, which were found by this review, are driven by the reporting domain. Missing statistical inference information might be one of the contributor to this negative association. Included articles successfully

delivered information on all reporting items except for the statistical inference. Only four randomized studies reported confidence intervals around ROI. Providing confidence intervals for ROI is not common because ROI is measured as a ratio. An additional method such as bootstrapping is needed. This could be the easiest improvement in WWP evaluations.

Second, having a positive effect for the internal validity is the most important result from a policy perspective. Recent critiques of the WWP ROI literature suggest that studies with greater internal validity should be yielding lower ROI estimates. That is, according to critics, ROI is expected to be lower with higher rigor in the methods that control for selection, deal with data issues, and detailed program costs. Yet we find that studies with greater internal validity (i.e., with stronger evidence for causal inference) have higher ROI estimates. In general, evaluation studies, regardless of study design, do not provide the distribution information of the benefits including outliers, which could be one of the contributor for positive association in estimation methods and ROI findings.

In addition to methodological quality, this paper examines critiques that are associated with effectiveness of WWPs, such as organizations' characteristics, program content, and conflict of interest. Industry, size, workplace environment, and resources for wellness programs are essential workplace characteristics that have an impact on reasons for WWP adoption and evaluation. Despite the popularity of these programs, the peer-reviewed papers showed that the majority of evaluations have been done in large organizations with the motivation of cost savings and positive ROI. WWP evaluations need to include clear information about all organizations investing in WWPs, especially

about their motivation for adopting them. Even though the ACA has directed more resources to organizations for WWP evaluations, the issue remains. There is more information about how high cost health plans can be modified using wellness programs for large organizations, but not enough proof for small organizations and information on organizations' reason for adoption other than positive ROI.

One other issue is the content itself of WWPs. The findings of this paper show that WWPS with a disease management component report higher ROI. We need to be able to identify whether the evaluated WWP include a disease management or health prevention component. These components target a specific health condition, whereas wellness components target health behavior. Thus, WWPs with a specific outcome target could save more money compared to WWPs with only general wellness or health behavior targets.

Finally, analyses showed that conflicts of interest arose due to evaluating the program internally, which resulted in higher ROI. Independent evaluation is an essential element to increasing the studies rigor. Eliminating conflicts of interest may be one of the hardest obstacles in the field due to the need to rely on the cooperation of the WWP host organizations.

This systematic review has two main limitations. First, this study focused on ROI findings due to the focus of economic evaluation literature and the critiques in the field. However, the relevant outcomes from employers' perspectives are varied and subject to change based on organizations' characteristics. For example, a small nonprofit organization in a specific industry might adopt a WWP for corporate citizenship

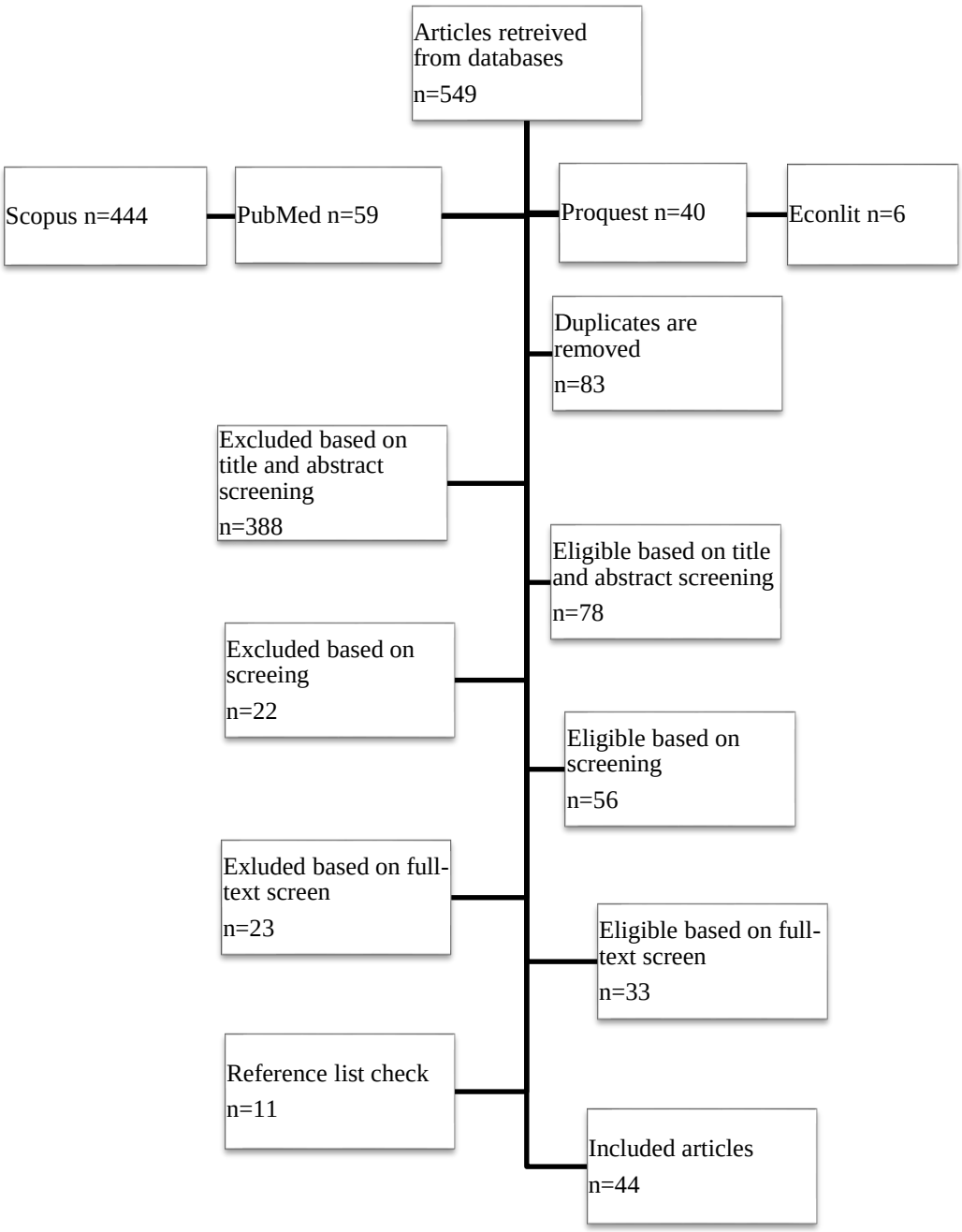
purposes, whereas a large for-profit organization in the same industry might adopt a WWP to reduce turnover rate. The field clearly has neglected this distinction. Second, although the field can provide improvements in methods to deal with measurement errors in data collection and self-selection into program participation, these inherent limitations cannot be eliminated. Randomized clinical trials are very difficult, if not impossible, for legal and logistical reasons. In addition, some of the health-related data are not available even to an independent evaluator due to HIPAA regulations.

This paper advances the field of economic evaluations of workplace wellness programs, in general, by providing information on areas that can improve methodological quality. Critiques in this field suggested that studies with higher internal validity resulted in lower ROI and the literature needed methods to improve internal validity, such as dealing with selection issues and estimation methods. However, this paper showed that the negative relationship between methodological quality and ROI was significant due to low reporting quality of papers. Lack of statistical inference information around ROI is an important issue. We cannot conduct a meta-analysis to derive common effects of WWPs when statistical inference information is missing. In addition, small organizations are underrepresented in this field. Although there are policies supporting small businesses adopting WWPs, we do not see that the field has the same focus on small organizations as much as larger companies and policymakers. The economic evaluation literature needs better reported peer-reviewed studies and attention on WWPs in companies with different characteristics, especially small companies with various and different reasons for WWP adoption. The research needs to validate whether WWPs can significantly impact public

health. The advancements suggested in this paper will help us understand private motivations for adoption decisions to align private and public motivations to receive policy support.

Figures

Figure 2.1. PRISMA Flow Diagram for Article Inclusion



Tables

Table 2.1. Search Strategy

Database	Search field	A. Participant		B. Intervention		C. Economic evaluation		D. Economic outcome	Final Search with Filters
PubMed	Abstract, title, and keywords	workplace OR employee OR worksite OR worker OR "work place" OR "work site"	AND	wellness OR health OR "health promotion" OR "health prevention" OR “well-being" OR wellbeing	AND	"cost benefit " OR "cost effectiveness" OR "cost analysis" OR "economic evaluation" OR "economic outcome" OR "economic analysis" OR "economic assessment"	AND	"return on investment" OR ROI	A AND B AND C AND D with filters English language and journal article
EconLit	Abstract, title, and keywords	workplace OR employee OR worksite OR worker OR "work place" OR "work site"	AND	wellness OR health OR "health promotion" OR "health prevention" OR “well-being" OR wellbeing	AND	"cost benefit " OR "cost effectiveness" OR "cost analysis" OR "economic evaluation" OR "economic outcome" OR "economic analysis" OR "economic assessment"	AND	"return on investment" OR ROI	A AND B AND C AND D with filters English language, peer reviewed articles, and scholarly journals

Table 2.1. Search Strategy (Continued)

Database	Search field	A. Participant		B. Intervention		C. Economic evaluation		D. Economic outcome	Final Search with Filters
Proquest Central	Abstract, title, and subject	workplace OR employee OR worksite OR worker OR "work place" OR "work site"	AND	wellness OR health OR "health promotion" OR "health prevention" OR “well-being" OR wellbeing	AND	"cost benefit " OR "cost effectiveness" OR "cost analysis" OR "economic evaluation" OR "economic outcome" OR "economic analysis" OR "economic assessment"	AND	"return on investment" OR ROI	A AND B AND C AND D with filters English language, peer reviewed articles, and scholarly journals
Scopus	Abstract, title, and keywords	workplace OR employee OR worksite OR worker OR "work place" OR "work site"	AND	wellness OR health OR "health promotion" OR "health prevention" OR “well-being" OR wellbeing	AND	"cost benefit " OR "cost effectiveness" OR "cost analysis" OR "economic evaluation" OR "economic outcome" OR "economic analysis" OR "economic assessment"	AND	"return on investment" OR ROI	A AND B AND C AND D with filters English, journal, and article

Table 2.2. Detailed Information for Included Articles

Article	Program Content	Country	Company/ Industry	Size of the Company	Participants (Non-participants)	Follow-up Length (Years)	Evaluation Start Year
Randomized studies (N=13)							
Barbosa C. et al., 2015[5]	Work and family (Wellness program)	United States	Information technology firm	Large	473 (473)	1.5	
Oude Hengel K. M. et al., 2014[40]	Prevention program (Wellness program)	Netherlands	Six construction companies (house, commercial or industrial building	N/A	170 (119)	1	
van Dongen J. M. et al., 2016[54]	Mindfulness training (Wellness program)	Netherlands	Two Dutch governmental institute	Large	129 (128)	1	2010
Groeneveld I.F. et al., 2011[19]	Lifestyle intervention (Wellness program)	Netherlands	Construction indsutry-15 plants	Large	293 (280)	1	
van Dongen J. et al., 2013[53]	Worksite vitality intervention (Wellness program)	Netherlands	Two Dutch academic hospital:	Large	367 (363)	2	2009
Noben C. et al., 2015[36]	Mental health program (Wellness program)	Netherlands	A Dutch Academic Hospital	Large	207 (206)	0.5	
Proper K. I. et al., 2004	Worksite physical activity counselling (Wellness program)	Netherlands	Three municipal services of a Dutch town	Large	97 (167)	3	1999
van Holland et al., 2017[55]	Worker health surveillance program (Wellness program)	Netherlands	Dutch meat processing company	Large	303 (683)	3	2012
Robroek S. J. W. et al., 2012[45]	Workplace health program (Wellness program)	Netherlands	Two health care organizations, 2 commercial services, 2 executive government branches	Large	465 (459)	2	2010

Table 2.2. Detailed Information for Included Articles (Continued)

Article	Program Content	Country	Company/ Industry	Size of the Company	Participants (Non-participants)	Follow-up Length (Years)	Evaluation Start Year
Randomized studies (N=13)							
Robroek S. J. W. et al., 2012[45]	Workplace health program (Wellness program)	Netherlands	Two health care organizations, 2 commercial services, 2 executive government branches	Large	465 (459)	2	2010
Steinberg G. et al., 2015[50]	Personalized wellness program (Wellness program)	United States	Aetna	Large	264 (945)	2	2013
Meenan R. T. et al., 2010[30]	Obesity prevention (Wellness program)	United States	Oahu, Hawaii hotel workers	Large	3346	3	2006
Milani R. V. et al., 2009[33]	Workplace wellness intervention (Disease management)	United States	A single employer, 2 locations (1 cite treatment the other is control)	N/A	185 (154)	1	
Thiart H. et al., 2016[51]	Insomnia therapy (Disease management)	Germany	School teachers	N/A	64 (64)	0.5	2013
Quasi-experimental studies (N=4)							
Ozminkowski R. J. et al., 1999[41]	Health management program (Wellness program)	United States	Citibank	Large	11194 (11644)	3	1994
Grossmeier J. et al., 2013[20]	Health management program (Disease management and wellness programs)	United States	BP America US employees	Large	29642 (32825)	3	2009
Musich S. et al., 2015[34]	Health management program (Disease management and wellness programs)	United States	Dell	Large	12037 (12614)	4	2009

Table 2.2. Detailed Information for Included Articles (Continued)

Article	Program Content	Country	Company/ Industry	Size of the Company	Participants (Non-participants)	Follow-up Length (Years)	Evaluation Start Year
Serxner S. et al., 2012[49]	Health management (Disease management and wellness programs)	United States	A large Financial Services Corporation	Large	28818 (8574)	5	2003
Observational cohort studies with control group (N=3)							
Light E. M. W. et al., 2015[25]	Workplace wellness program (Disease management and wellness programs)	United States	Price Cooper, Golub corporation (large retail grocery org)	Large	879 (879)	5	2007
Michaud T. L. et al., 2016[32]	Health promotion program (Wellness Program)	United States	University of Minnesota	Large	1501 (86389)	3	2010
Liu H. et al., 2013[26]	Health & wellness (Disease management and wellness programs)	United States	PepsiCo	Large	24503 (30525)	5	2002
Observational case studies with control group (N=16)							
Naydeck B. L. et al., 2008[35]	Wellness program	United States	Highmark	Large	1892 (1892)	4	2002
Yen L. et al., 2010[56]	Health promotion program (Wellness program)	United States	Midwest utility company	Large	2036 (717)	8	2000
Bertera, R. L., 1990[6]	Workplace health program (Wellness program)	United States	large multi-location (60 sites) diversified industrial company	Large	29315 (14573)	3	1984
Griffin S. C. et al., 2016[18]	Fitness (Wellness Program)	United States	Tucson fire department	Small	32 (77)	4	2007
Goetzel et al., 1998[14]	Workplace health program (Wellness program)	United States	Cincinnati headquarters of The Procter & Gamble Company	Large	3993 (4341)	3	1990
Schultz A. B. et al., 2002[47]	Workplace health promotion program (Wellness program)	United States	Two manufacturing plants in the Midwest	Large	2596 (1593)	6	1995

Table 2.2. Detailed Information for Included Articles (Continued)

Article	Program Content	Country	Company/ Industry	Size of the Company	Participants (Non-participants)	Follow-up Length (Years)	Evaluation Start Year
Serxner S. et al., 2003[48]	Health promotion program (Wellness Program)	United States	DaimlerChrysler AG 14 worksite	Large	13048 (13363)	1	1997
Abraham J.M. et al., 2012[1]	Fitness program (Wellness program)	United States	University of Minnesota	Large	194	3	2006
Jutkowitz E. et al., 2015[23]	Disease management program (Disease management and wellness programs)	United States	University of Minnesota	Large	3746 (7585)	6	2004
Nyman J.A. et al., 2013[39]	Disease management program (Disease management)	United States	University of Minnesota	Large	4226 (8677)	7	2004
Nyman J.A. et al., 2012[38]	Health promotion program (Disease management and wellness program)	United States	University of Minnesota	Large	2073 (4072)	5	2004
Kapinos K. A. et al., 2015[24]	Workplace wellness program (Disease management and wellness programs)	United States	A large firm	Large	1431 (1264)	9	2003
Liu H. et al., 2013[27]	Wellness program (Disease management and wellness programs)	United States	PepsiCo	Large	6623 (6623)	6	2002
Nyman J.A. et al., 2010[37]	Health promotion (Disease management and wellness programs)	United States	University of Minnesota	Large	3619 (1757)	2	2006
Caloyeras J.P. et al., 2014[9]	Wellness program (Disease management and wellness programs)	United States	PepsiCo	Large	22880 (22204)	10	2002
Mattke S. et al., 2009[29]	Population health management program (Disease management and wellness programs)		Two large employers in the consumer goods industry	Large	39809 (158962)	5	2001

Table 2.2. Detailed Information for Included Articles (Continued)

Article	Program Content	Country	Company/ Industry	Size of the Company	Participants (Non-participants)	Follow-up Length (Years)	Evaluation Start Year
Observational studies without control group (N=8)							
Bowne D. et al., 1984[8]	Industrial fitness program (Wellness program)	United States	Southwestern home office / Houston	Large	184	5	1977
Iijima S. et al., 2013[22]	Mental health (Wellness program)	Japan	Eleven major companies (6 wholesale dealers, 3 transportation, and 2 production companies	Large	1169	2	2011
Golaszewski T. et al., 1992[17]	Health promotion program (Wellness Program)	United States	Travelers Insurance company	Large	39809	15	1986
Palumbo M.V. et al., 2013[43]	Health prevention (Wellness program)	United States	One hospital unit	Small	48	5.5	2008
Ozminkowski R. J. et al., 2002[42]	Health and wellness program (Disease management and wellness programs)	United States	Johnson and Johnson	Large	18331	5	1990
Maniscalco P. et al., 1999[28]	Wellness program	United States	The Lafayette Offshore Business Unit/ Louisiana	Small	91	5	1993
Baker K. M. et al., 2008[4]	Obesity management (Disease management)	United States	American Specialty Health, INC (ASH) multiple workplaces (119 companies)	Large	890	1	2006
Bevis C. C. et al., 2014[7]	Wellness program (Disease management and wellness programs)	United States	Major employer in Orlando area	Large	224	3	2006

Table 2.3. Summary of Characteristics for Included Articles

Article Characteristics (N=44)	**Frequency**	**Percentage**
Randomized studies	13	29.55
Quasi-experimental studies	4	6.82
Observational cohort studies with control group	3	6.82
Observational case studies with control group	16	36.36
Observational studies without control group	8	18.18
Evaluated program: Only wellness	28	63.64
Evaluated program: Only disease management	5	11.36
Evaluated program: Wellness and disease management	11	25.00
Country of evaluated program: United States	34	77.27
Country of evaluated program: Netherlands	8	19.51
Country of evaluated program: Germany	1	2.44
Country of evaluated program: Japan	1	2.44
Program in a small company	3	6.82
Article provides ROI findings	28	63.64

Table 2.4. Outcome, ROI Formula, ROI, and Quality Index Information for ROI Analysis Sample

Article	Outcome Details	ROI formula in article (Unit)	Provided ROI	Recalculated ROI	Statistical Inference	Quality Index
Randomized studies (N=11)						
Barbosa C. et al., 2015[5]	Health care utilization (Inpatient, outpatient, and ER) presenteeism, and turnover	(Benefit-Cost)/ Cost (Ratio)	1.68	1.68	Yes	17.00
Oude Hengel K. M. et al., 2014[40]	Absenteeism and presenteeism	[(Benefit-Cost)/ Cost] *100 (percentage)	543	5.43	No	14.67
van Dongen J. M. et al., 2016[54]	Costs of medical and occupational health, absenteeism, and presenteeism	(Benefit-Cost)/ Cost (Ratio)	-3.51	-3.52	Yes	14.67
Groeneveld I.F. et al., 2011[19]	Absenteeism			-0.49	Yes	14.33
van Dongen J. et al., 2013[53]	Health care costs (General practitioner, complementary medicine, medical specialist, and hospitalization), absenteeism and presenteeism	Benefit/ Cost (Ratio)	-2.21	-3.83	No	14.00
Noben C. et al., 2015[36]	Absenteeism and presenteeism	Benefit/ Cost (Ratio)	11	10.17	No	13.67
van Holland et al., 2017[55]	Absenteeism and presenteeism	(Benefit-Cost)/ Cost (Ratio)	-11.6	-11.61	Yes	13.00
Proper K. I. et al., 2004	Cost of sick leave			0.19	Yes	13.00
Meenan R. T. et al., 2010[30]	Health care costs, productivity, and absenteeism			-0.98	No	11.67
Milani R. V. et al., 2009[33]	Health care costs		6		No	8.33

Table 2.4. Outcome, ROI Formula, ROI, and Quality Index Information for ROI Analysis Sample (Continued)

Article	Outcome Details	ROI formula in article (Unit)	Provided ROI	Recalculated ROI	Statistical Inference	Quality Index
Thiart H. et al., 2016[51]	Absenteeism and presenteeism	[(Benefit-Cost)/ Cost] *100 (percentage)	208.81	2.09	Yes	13.33
Quasi-experimental studies (N=4)						
Ozminkowski R. J. et al., 1999[41]	Health care costs	Benefit/ Cost (Ratio)	4.73	3.73	No	11.58
Grossmeier J. et al., 2013[20]	Health care costs (Inpatient, office visits, and ER)	Benefit/ Cost (Ratio)	3	2.00	No	12.08
Musich S. et al., 2015[34]	Health care costs (Medical, pharmaceutical, and short-term disability) and productivity	Benefit/ Cost (Ratio)	2.48	1.48	No	11.58
Serxner S. et al., 2012[49]	Health care costs (Medical and prescription claims)	Benefit/ Cost (Ratio)	2.45	1.45	No	11.08
Observational cohort studies with control group (N=1)						
Light E. M. W. et al., 2015[25]	Medical claims (Inpatient, outpatient, pharmaceutical, and professional), risk levels, and presence of comorbidities	Benefit/ Cost (Ratio)	4.33	3.33	No	12.00
Observational case studies with control group (N=9)						
Naydeck B. L. et al., 2008[35]	Medical claims (Inpatient, outpatient, pharmaceutical, and professional)	Benefit/ Cost (Ratio)	1.65	0.65	No	10.75
Yen L. et al., 2010[56]	Health care costs and time away from work	Benefit/ Cost (Ratio)	1.66	0.66	No	10.25
Bertera, R. L., 1990[6]	Absenteeism (Disability days) and employment cost	Benefit/ Cost (Ratio)	1.42	0.42	No	9.75
Griffin S. C. et al., 2016[18]	Injury and worker's comp claims	[(Benefit-Cost)/ Cost] *100 (percentage)	2.4	0.02	No	9.42

Table 2.4. Outcome, ROI Formula, ROI, and Quality Index Information for ROI Analysis Sample (Continued)

Article	Outcome Details	ROI formula in article (Unit)	Provided ROI	Recalculated ROI	Statistical Inference	Quality Index
Schultz A. B. et al., 2002[47]	Disability days	(Benefit-Cost)/ Cost (Ratio)	1.3	1.27	No	8.92
Nyman J.A. et al., 2013[39]	Hospitalization, avoidable hospitalization and costs (Inpatient, physician care, lab and pathology claims, pharmacy, radiology, surgery, and ER)	Benefit/ Cost (Ratio)	1.63	0.63	No	10.75
Nyman J.A. et al., 2012[38]	Health care costs and absenteeism	Benefit/ Cost (Ratio)	1.76	0.76	No	11.58
Nyman J.A. et al., 2010[37]	Health care costs and absenteeism	Benefit-Cost (USD)	-625,947	-0.13	No	10.42
Caloyeras J.P. et al., 2014[9]	Health care costs (Hospital admissions) and absenteeism	no information	1.46		No	10.08
Observational studies without control group (N=6)						
Bowne D. et al., 1984[8]	Disability and health care cost	(Benefit-Cost)/ Cost (Ratio)	1.93	1.93	No	9.33
Iijima S. et al., 2013[22]	Labor cost of mental health (Absenteeism)	Benefit/ Cost (Ratio)	1.55	0.55	No	7.67
Palumbo M.V. et al., 2013[43]	Hours of unscheduled absence as a proxy to wellness	Benefit-Cost (USD)	3,747	0.49	No	7.67
Golaszewski T. et al., 1992[17]	Health care costs, productivity, absenteeism, and life insurance claims	(Benefit-Cost)/ Cost (Ratio)	2.4	2.43	No	7.67
Maniscalco P. et al., 1999[28]	Health care costs and productivity	Benefit/ Cost (Ratio)	2.51	1.51	No	5.00
Baker K. M. et al., 2008[4]	Health care costs (Inpatient, outpatient, and pharmaceutical) and presenteeism	Benefit/ Cost (Ratio)	1.17	0.17	No	9.33

Table 2.5. ROI Content and Formulation for Articles that Provided ROI

Formulation for ROI (N=28)	**Frequency**	**Percentage**
Net benefit to cost ratio	6	21.43
Net benefit to cost as percentage	3	10.71
Net benefit	2	7.14
Benefit to cost ratio	15	53.57
No information	2	7.14
Statistical inference information on ROI (N=28)		
Any statistical information is provided	4	14.81

Table 2.6. Summary of Quality Indices, Checklist Domains, and Recalculated ROI for Analysis Sample

Analysis Outcome (N=29)	**Mean**	**Standard deviation**	**Min**	**Max**
Recalculated ROI	0.68	3.57	-12.61	10.17
Quality Index (N=29)				
Items are equally weighted	11.97	2.95	5.00	18.00
Reduced quality index	11.66	2.57	5.00	16.00
Rigor Rubric Main Domains (N=29)				
Reporting (8 items)	6.92	0.75	4.00	8.00
Internal Validity (7 items)	4.74	2.05	1.00	8.00
External validity (1 item)	0.21	0.41	0.00	1.00
Power (1 item)	0.10	0.31	0.00	1.00

Table 2.7. Summary of Characteristics for Analysis Sample That Has Recalculated ROI Measure

Article Characteristics (N=29)		
Program in a small company	3	10.34
Evaluated program: Only wellness	20	68.97
Evaluated program: Only disease management	4	13.79
Evaluated program: Wellness and disease management	5	17.24
ROI with only health care cost	7	24.14
ROI with only absenteeism cost	7	24.14
ROI with health care and absenteeism	4	13.79
ROI with health care and productivity	4	13.79
ROI with absenteeism and productivity	3	10.34
ROI with health care, absenteeism, and productivity	4	13.79
Published after 2010	17	58.62
Study duration is at least three years	18	62.07
Internal validity item 1: Study design (N=29)	**Frequency**	**Percentage**
Randomized studies	10	34.48
Quasi-experimental studies	4	13.79
Observational cohort studies with control group	1	3.45
Observational case studies with control group	8	27.59
Observational studies without control group	6	20.69
Internal validity item 3: Estimation method (N=29)		
Appropriate methods for outcome estimates	7	24.14
Internal validity item 4: Valuation of cost (N=29)		
Program cost measures are clearly described	9	31.03
Conflict of Interest (N=29)		
Based on funding source	9	32.14

Table 2.8. Mean Effects on Recalculated ROI: Quality Index (Items Equally Weighted) and Article Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Quality index: Items are equally weighted	-0.033						-0.018
	(0.233)						(0.405)
Programs in a small company[1]		-0.010					0.094
		(2.220)					(3.462)
Only disease management program[2`]			1.272				0.434
			(2.013)				(2.620)
Disease management and wellness programs[2]			0.729				1.052
			(1.838)				(2.473)
Only absenteeism costs[3]				-0.564			-1.771
				(1.969)			(2.923)
Health care and absenteeism costs[3]				-0.884			-2.293
				(2.309)			(3.155)
Any productivity cost[3,4]				-1.968			-3.163
				(1.781)			(2.421)
Publication year dummy (2011)[5]					-0.280		-1.485
					(1.372)		(2.200)
Evaluation duration (at least 3 years)[6]						-0.747	-2.468
						(1.386)	(2.198)
Constant	1.084	0.684	0.382	1.688	0.848	1.147	4.993
	(2.871)	(0.714)	(0.822)	(1.392)	(1.050)	(1.092)	(6.452)
Observations	29	29	29	29	29	29	29

Notes: 1. Small is 1 if the company has 500 and less employees, 0 otherwise.
2. The base category is "Only wellness program"
3. The base category is only health care cost.
4. This component includes combinations of costs that include productivity (presenteeism and else)
5. Year dummy is 1 if the study was published after 2010 (2011 and after)
6. Evaluation duration is 1 if the study period is at least 3 years

Table 2.9. Mean Effects on Recalculated ROI: Reduced Quality Index and Quality Index Domains and Article Characteristics

	(1)	(2)	(3)	(4)
Reduce quality index	-0.032	-0.041		
	(0.267)	(0.455)		
Reporting (8 items)			-2.474**	-3.320**
			(1.061)	(1.346)
Internal validity (8 items)			0.672*	0.753
			(0.387)	(0.507)
Programs in a small company[1]		-0.017		-1.825
		(3.485)		(3.150)
Only disease management program[2]		0.428		0.277
		(2.584)		(2.276)
Disease management and wellness programs[2]		1.049		-0.509
		(2.447)		(2.239)
Only absenteeism costs[3]		-1.803		-0.587
		(2.929)		(2.623)
Health care and absenteeism costs[3]		-2.313		-1.338
		(3.157)		(2.806)
Any productivity cost[3,4]		-3.178		-3.515
		(2.428)		(2.142)
Publication year dummy (2011)[5]		-1.447		-0.371
		(2.171)		(1.958)
Evaluation duration (at least 3 years)[6]		-2.477		-2.106
		(2.169)		(1.915)
Constant	1.059	5.271	14.616**	23.508**
	(3.188)	(6.766)	(6.444)	(9.306)
Observations	29	29	29	29

Notes: 1. Small is 1 if the company has 500 and less employees, 0 otherwise.
2. The base category is "Only wellness program"
3. The base category is only health care cost.
4. This component includes combinations of costs that include productivity (presenteeism and else)
5. Year dummy is 1 if the study was published after 2010 (2011 and after)
6. Evaluation duration is 1 if the study period is at least 3 years

Table 2.10. Mean Effects on Calculated ROI: Quality Index Some Subcomponents

	(1)	(2)	(3)	(4)	(5)
Observational studies with control group	-0.053			-0.785	
	(1.946)			(2.225)	
Quasi-experimental	1.265			1.139	
	(2.384)			(2.599)	
Randomized	-1.086			-2.226	
	(1.907)			(2.227)	
Appropriate methods for outcome estimates		0.736		0.251	
		(1.573)		(1.853)	
Program cost measures are clearly described			0.559	1.859	
			(1.457)	(1.794)	
Conflict of interest					0.876
					(1.490)
Constant	0.900	0.506	0.510	0.900	0.273
	(1.508)	(0.773)	(0.812)	(1.536)	(0.845)
Observations	29	29	29	29	28

References

1. Abraham, J. M., Nyman, J. A., Feldman, R., & Barleen, N. (2012). The effect of participation in a fitness rewards program on medical care expenditures in an employee population. *Journal of Occupational and Environmental Medicine*, *54*(3), 280-285.

2. Adams, M. E., McCall, N. T., Gray, D. T., Orza, M. J., & Chalmers, T. C. (1992). Economic analysis in randomized control trials. *Medical Care*, 231-243.

3. Aldana, S. G. (2001). Financial impact of health promotion programs: A comprehensive review of the literature. *American Journal of Health Promotion*, 15(5): 296–320.

4. Baker K.M., Goetzel R.Z., Pei X., Weiss A.J., Bowen J., Tabrizi M.J., Nelson C.F., Metz R.D., Pelletier K.R., Thompson E. (2008). Using a return-on-investment estimation model to evaluate outcomes from an obesity management worksite health promotion program. *Journal of Occupational and Environmental Medicine*, *50*(9), 981-990.

5. Barbosa C., Bray J.W., Dowd W.N., Mills M.J., Moen P., Wipfli B., Olson R., Kelly E.L. (2015). Return on investment of a work–family intervention: Evidence from the work, family, and health network. *Journal of Occupational and Environmental Medicine, 57*(9), 943-951.

6. Bertera, R. L. (1990). The effects of workplace health promotion on absenteeism and employment costs in a large industrial population. *American Journal of Public Health*, *80*(9), 1101-1105.

7. Bevis C.C., Nogle J.M., Forges B., Chen P.C., Sievers D., Lucas K.R., Mahoney J.J., Crawford J.M. (2014). Diabetes wellness care: a successful employer-endorsed program for employees. *Journal of Occupational and Environmental Medicine*, *56*(10), 1052-1061.

8. Bowne, D. W., Russell, M. L., Morgan, J. L., Optenberg, S. A., & Clarke, A. E. (1984). Reduced Disability and Health Care Costs in an Industrial Fitness Program. *Journal of Occupational and Environmental Medicine, 26*(11), 809-816.

9. Caloyeras, J. P., Liu, H., Exum, E., Broderick, M., & Mattke, S. (2014). Managing manifest diseases, but not health risks, saved PepsiCo money over seven years. *Health Affairs*, *33*(1), 124-131.

10. Dawson, R. (2011). How significant is a boxplot outlier? *Journal of Statistics Education, 19*(2).

11. Dement, J. M., Epling, C., Joyner, J., & Cavanaugh, K. (2015). Impacts of workplace health promotion and wellness programs on health care utilization and costs: results from an academic workplace. *Journal of Occupational and Environmental Medicine*, 57(11), 1159-1169.

12. Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology and Community Health, 52*(6), 377-384.

13. Evers, S., Goossens, M., De Vet, H., Van Tulder, M., & Ament, A. (2005). Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria. *International Journal of Technology Assessment in Health Care*, 21(02), 240-245.

14. Gerard, K. (1992). Cost-utility in practice: a policy maker's guide to the state of the art. *Health Policy, 21*(3), 249-279.

15. Goetzel, R. Z., Jacobson, B. H., Aldana, S. G., Vardell, K., & Yee, L. (1998). Health care costs of worksite health promotion participants and non-participants. *Journal of Occupational and Environmental Medicine, 40*(4), 341-346.

16. Goetzel, R. Z., & Ozminkowski, R. J. (2008). The health and cost benefits of work site health-promotion programs. *Annual Review Public Health*, 29, 303-323.

17. Goetzel, R. Z., Tabrizi, M., Henke, R. M., Benevent, R., Brockbank, C. V. S., Stinson, K. & Newman, L. S. (2014). Estimating the return on investment from a health risk management program offered to small Colorado-based employers. *Journal of Occupational and Environmental Medicine*, 56(5), 554-560.

18. Golaszewski, T., Snow, D., Lynch, W., Yen, L., & Solomita, D. (1992). A Benefit-to-cost Analysis of a Work-site Health Promotion Program. *Journal of Occupational Medicine*, 34(12), 1164-1172.

19. Griffin, S. C., Regan, T. L., Harber, P., Lutz, E. A., Hu, C., Peate, W. F., & Burgess, J. L. (2016). Evaluation of a fitness intervention for new firefighters: injury reduction and economic benefits. *Injury Prevention*, 22(3), 181-188.

20. Groeneveld, I. F., van Wier, M. F., Proper, K. I., Bosmans, J. E., van Mechelen, W., & van der Beek, A. J. (2011). Cost-effectiveness and cost-benefit of a lifestyle intervention for workers in the construction industry at risk for cardiovascular disease. *Journal of Occupational and Environmental Medicine, 53*(6), 610-617.

21. Grossmeier, J., Seaverson, E. L., Mangen, D. J., Wright, S., Dalal, K., Phalen, C., & Gold, D. B. (2013). Impact of a comprehensive population health management program on health care costs. *Journal of Occupational and Environmental Medicine, 55*(6), 634-643.

22. Henke, R. M., Goetzel, R. Z., McHugh, J., & Isaac, F. (2011). Recent experience in health promotion at Johnson & Johnson: lower health spending, strong return on investment. *Health Affairs*, 30(3), 490-499.

23. Iijima, S., Yokoyama, K., Kitamura, F., Fukuda, T., & Inaba, R. (2013). Cost-benefit analysis of comprehensive mental health prevention programs in Japanese workplaces: a pilot study. *Industrial Health*, *51*(6), 627-633.

24. Jutkowitz, E., Nyman, J. A., Michaud, T. L., Abraham, J. M., & Dowd, B. (2015). For what illnesses is a disease management program most effective? J*ournal of Occupational and Environmental Medicine, 57*(2), 117-123.

25. Kapinos, K. A., Caloyeras, J. P., Liu, H., & Mattke, S. (2015). Does Targeting Higher Health Risk Employees or Increasing Intervention Intensity Yield Savings in a Workplace Wellness Program? *Journal of Occupational and Environmental Medicine, 57*(12), 1257-1261.

26. Light, E. M., Kline, A. S., Drosky, M. A., & Chapman, L. S. (2015). Economic analysis of the return-on-investment of a worksite wellness program for a large multistate retail grocery organization. *Journal of Occupational and Environmental Medicine,* 57(8), 882-892.

27. Liu, H., Harris, K. M., Weinberger, S., Serxner, S., Mattke, S., & Exum, E. (2013). Effect of an employer-sponsored health and wellness program on medical cost and utilization. *Population Health Management*, 16(1), 1-6.

28. Liu, H., Mattke, S., Harris, K. M., Weinberger, S., Serxner, S., Caloyeras, J. P., & Exum, E. (2013). Do Workplace Wellness Programs Reduce Medical Costs? Evidence from a Fortune 500 Company. *Inquiry*, 150-158.

29. Maniscalco, P., Lane, R., Welke, M., Mitchell, J. H., & Husting, L. (1999). Decreased rate of back injuries through a wellness program for offshore petroleum employees. *Journal of Occupational and Environmental Medicine, 41*(9), 813-820.

30. Mann, P. S. (2007). *Introductory statistics*. John Wiley & Sons.

31. Mattke, S., Serxner, S. A., Zakowski, S. L., Jain, A. K., & Gold, D. B. (2009). Impact of 2 employer-sponsored population health management programs on medical care cost and utilization. *The American Journal of Managed Care*, *15*(2), 113-120.

32. Meenan, R. T., Vogt, T. M., Williams, A. E., Stevens, V. J., Albright, C. L., & Nigg, C. (2010). Economic Evaluation of a Worksite Obesity Prevention and Intervention Trial Among Hotel Workers in Hawaii. *Journal of Occupational and Environmental Medicine, 52*(1S), S8-S13.

33. Merrill, R. M., Hyatt, B., Aldana, S. G., & Kinnersley, D. (2011). Lowering employee health care costs through the healthy lifestyle incentive program. *Journal of Public Health Management and Practice,* 17(3), 225-232.

34. Michaud, T. L., Nyman, J. A., Jutkowitz, E., Su, D., Dowd, B., & Abraham, J. M. (2016). Effect of workplace weight management on health care expenditures and quality of life. *Journal of Occupational and Environmental Medicine,* 58(11), 1073-1078.

35. Milani, R. V., & Lavie, C. J. (2009). Impact of worksite wellness intervention on cardiac risk factors and one-year health care costs. *The American Journal of Cardiology, 104*(10), 1389-1392.

36. Musich, S., McCalister, T., Wang, S., & Hawkins, K. (2015). An evaluation of the well at Dell health management program: health risk change and financial return on investment. *American Journal of Health Promotion, 29*(3), 147-157.

37. Naydeck, B. L., Pearson, J. A., Ozminkowski, R. J., Day, B. T., & Goetzel, R. Z. (2008). The impact of the Highmark employee wellness programs on 4-year healthcare costs. *Journal of Occupational and Environmental Medicine, 50*(2), 146-156.

38. Noben, C., Evers, S., Nieuwenhuijsen, K., Ketelaar, S., Gärtner, F., Sluiter, J., & Smit, F. (2015). Protecting and promoting mental health of nurses in the hospital setting: Is it cost-effective from an employer's perspective? *International journal of Occupational Medicine and Environmental Health*, *28*(5), 891.

39. Nyman, J. A., Barleen, N. A., & Abraham, J. M. (2010). The effectiveness of health promotion at the University of Minnesota: expenditures, absenteeism, and participation in specific programs. *Journal of Occupational and Environmental Medicine, 52*(3), 269-280.

40. Nyman, J. A., Abraham, J. M., Jeffery, M. M., & Barleen, N. A. (2012). The effectiveness of a health promotion program after 3 years: evidence from the University of Minnesota. *Medical Care, 50*(9), 772.

41. Nyman, J. A., Jeffery, M. M., Abraham, J. M., Jutkowitz, E., & Dowd, B. E. (2013). The 3-year disease management effect: understanding the positive return on investment. *Journal of Occupational and Environmental Medicine*, *55*(11), 1356-1364.

42. Oude Hengel, K. M., Bosmans, J. E., Van Dongen, J. M., Bongers, P. M., Van der Beek, A. J., & Blatter, B. M. (2014). Prevention program at construction worksites aimed at improving health and work ability is cost-saving to the employer: Results from an RCT. *American Journal of Industrial Medicine*, *57*(1), 56-68.

43. Ozminkowski, R. J., Dunn, R. L., Goetzel, R. Z., Cantor, R. I., Murnane, J., & Harrison, M. (1999). A return on investment evaluation of the Citibank, NA, Health Management Program. *American Journal of Health Promotion*, *14*(1), 31-43.

44. Ozminkowski, R. J., Ling, D., Goetzel, R. Z., Bruno, J. A., Rutter, K. R., Isaac, F., & Wang, S. (2002). Long-term impact of Johnson & Johnson's Health & Wellness Program on health care utilization and expenditures. *Journal of Occupational and Environmental Medicine*, *44*(1), 21-29.

45. Palumbo, M. V., Sikorski, E. A., & Liberty, B. C. (2013). Exploring the cost-effectiveness of unit-based health promotion activities for nurses. *Workplace Health & Safety*, *61*(12), 514-520.

46. Phillips, J. J., & Phillips, P. P. (2007). Show me the money: how to determine ROI in people, projects, and programs. Berrett-Koehler Publishers, Incorporated Retrieved from http://ebookcentral.proquest.com

47. Proper, K. I., De Bruyne, M. C., Hildebrandt, V. H., Van Der Beek, A. J., Meerding, W. J., & Van Mechelen, W. (2004). Costs, benefits and effectiveness of worksite physical activity counseling from the employer's perspective. *Scandinavian Journal of Work, Environment & Health*, 36-46.

48. Robroek, S. J., Polinder, S., Bredt, F. J., & Burdorf, A. (2012). Cost-effectiveness of a long-term Internet-delivered worksite health promotion programme on physical activity and nutrition: a cluster randomized controlled trial. *Health Education Research*, *27*(3), 399-410.

49. Sacristán, J. A., Soto, J., & Galende, I. (1993). Evaluation of pharmacoeconomic studies: utilization of a checklist. *Annals of Pharmacotherapy*, *27*(9), 1126-1133.

50. Schultz, A. B., Lu, C., Barnett, T. E., Yen, L. T. C., McDonald, T., Hirschland, D., & Edington, D. W. (2002). Influence of participation in a worksite health-promotion program on disability days. *Journal of Occupational and Environmental Medicine*, *44*(8), 776-780.

51. Serxner, S. A., Gold, D. B., Grossmeier, J. J., & Anderson, D. R. (2003). The Relationship Between Health Promotion Program Participation and Medical Costs: A Dose Response. *Journal of Occupational and Environmental Medicine*, *45*(11), 1196-1200.

52. Serxner, S., Alberti, A., & Weinberger, S. (2012). Medical cost savings for participants and nonparticipants in health risk assessments, lifestyle management, disease management, depression management, and nurseline in a large financial services corporation. *American Journal of Health Promotion, 26*(4), 245-252.

53. Steinberg, G., Scott, A., Honcz, J., Spettell, C., & Pradhan, S. (2015). Reducing metabolic syndrome risk using a personalized wellness program. *Journal of Occupational and Environmental Medicine, 57*(12), 1269-1274.

54. Thiart H., Ebert D.D., Lehr D., Nobis S., Buntrock C., Berking M., Smit F., Riper H. (2016). Internet-based cognitive behavioral therapy for insomnia: a health economic evaluation. *Sleep, 39*(10), 1769-1778.

55. Trogdon, J., Finkelstein, E. A., Reyes, M., & Dietz, W. H. (2009). A return-on-investment simulation model of workplace obesity interventions. *Journal of Occupational and Environmental Medicine,* 51(7), 751-758.

56. van Dongen, J. M., Strijk, J. E., Proper, K. I., van Wier, M. F., van Mechelen, W., van Tulder, M. W., & Van Der Beek, A. J. (2013). A cost-effectiveness and return-on-investment analysis of a worksite vitality intervention among older hospital workers: results of a randomized controlled trial. *Journal of Occupational and Environmental Medicine, 55*(3), 337-346.

57. van Dongen J.M., Van Berkel J., Boot C.R.L., Bosmans J.E., Proper K.I., Bongers P.M., Van Der Beek A.J., Van Tulder M.W., Van Wier M.F. (2016). Long-term cost-effectiveness and return-on-investment of a mindfulness-based worksite intervention: results of a randomized controlled trial. *Journal of Occupational and Environmental Medicine, 58*(6), 550-560.

58. van Holland, B. J., Reneman, M. F., Soer, R., Brouwer, S., & de Boer, M. R. (2017). Effectiveness and Cost-benefit Evaluation of a Comprehensive Workers' Health Surveillance Program for Sustainable Employability of Meat Processing Workers. *Journal of Occupational Rehabilitation*.

59. Yen, L., Schultz, A. B., Schaefer, C., Bloomberg, S., & Edington, D. W. (2010). Long-term return on investment of an employee health enhancement program at a Midwest utility company from 1999 to 2007. *International Journal of Workplace Health Management, 3*(2), 79-96.

Appendix 2.A. Additional Tables

Table 2.A1. Methodological Rigor Rubric and Quality Index

Domains	Checklist	Explanation	Responses	Source (Item Number)
1. Article	1.1. Who is the author?	Full list of authors		
	1.2. What is the title of the article?	Title of the article		
	1.3. What was year the study conducted/published?	Year that the study was published		
	1.4. If published which journal?	Journal that the study was published		
	1.5. What is evaluated?	Disease management (DM) or wellness program (WP)	DM or WP	Gerard, 1992 (2,3)
	1.6. What is the country of study?	Name of the country that the study was conducted		Gerard, 1992 (5)
	1.7. What is the industry/company?	Name of the industry or company		
	1.8. The size of the worksite that wellness program took place.	If 500 and less employees small company, and large otherwise	Small or Large	
	1.8.1 Number of participant and nonparticipants	Number of participant and nonparticipants an		
	1.9. What academic department or research center conducted the study?	To compare with funding agency (1.10 and 1.11) for possible conflicts of interest		Gerard, 1992 (6)
	1.10. Who is the funding agency?	To compare with academic department (1.9) for possible conflicts of interest		Gerard, 1992 (8)
	1.11. Are there authorial conflicting interests	Article states any conflicts of interest	0 or 1	Evers, 2005 (18)
Rigor				
2. Reporting	Are the followings items clearly described / reported before the results?	2.3, 2.6 and 2.8 are not necessarily before the results section because the reporting varies based on journals.		
	2.1. Objectives of the study	Objectives of the study are reported before results	0 or 1	Downs and Black, 1998 (1)
	2.2. Intervention(s)	Programs that are to be evaluated are clearly described before the results	0 or 1	Downs and Black, 1998 (4)

Table 2.A1. Methodological Rigor Rubric and Quality Index (Continued)

Domains	Checklist	Explanation	Responses	Source (Item Number)
	2.3. Study sample	The characteristics of the (non)participants included in the study are clearly described. Each sub question has equal weight in contributing to 2.3	0, 0.33, 0.67, or 1	Downs and Black, 1998 (3)
	2.3.1. Study population	The information on eligible employee population is provided	0 or 1	Evers, 2005 (1)
	2.3.2. Inclusion/exclusion criteria	The information on selection criteria from eligibility to participation is provided	0 or 1	Downs and Black, 1998 (3)
	2.3.3. Analysis sample	The information on analysis sample is provided	0 or 1	Downs and Black, 1998 (3)
	2.4. Type of the economic analysis	Type of the economic analysis (Cost-benefit, cost-effectiveness, cost-utility, cost analyses) is reported before results	0 or 1	Sacristan, 1993 (7)
	2.5. Main outcome(s)	Main outcomes including clinical and monetized outcomes are reported before results	0 or 1	Downs and Black, 1998 (2)
	2.6. Intervention costs	Overall program cost is reported. Details not needed for this question. Details are in 3.3 and 3.4	0 or 1	
	2.7. Main finding(s)	Main findings are reported before results. There is no need for statistical inference information for this question	0 or 1	Downs and Black, 1998 (6)
	2.8. Statistical inference information about the main outcomes (interquartile change, standard errors, standard deviations, confidence interval, p-values)	Statistical inference information is provided for all the outcomes	0 or 1	Downs and Black, 1998 (7)

Table 2.A1. Methodological Rigor Rubric and Quality Index (Continued)

Domains	Checklist	Explanation	Responses	Source (Item Number)
3. Internal validity				
	3.1. Were the main outcome measures used clearly described?	Clear description of what the outcomes included and how they were measured	0 or 1	Evers, 2005 (10,11,12)
	3.2. Were study subjects randomized to intervention groups? (0. Observational case or cohort without control group, 1. Observational case with control group, 2. Observational cohort with control group, 3. quasi-experimental, 4. Randomized	The rank range (0-4) is normalized to 0-1 range by weighting the answers.	0, 0.25, 0.50, 0.75, or 1	Downs and Black, 1998 (23)
	3.3. Was the method used to assess the main outcomes appropriate?	Average of 3.3.1.-3.3.2. Each sub question has equal weight in contributing to 3.3	0, 0.5, or 1	Downs and Black, 1998 (18)
	3.3.1. When not randomized, was any method used to deal with selection bias?	Whether the analyses on participants were intention-to-treat; whether the distribution of confounders in the different participant groups was described and differences were taken into account e.g. Matching method	0 or 1	Downs and Black, 1998 (25)
	3.3.2. Appropriate method for outcome estimates	The statistical techniques must be appropriate to the data such as if the data were not normally distributed (skewed data), or if nonparametric approaches were needed. Does the estimation method take skewness in the data into account (e.g. Count data, binary data models etc.)	0 or 1	

Table 2.A1. Methodological Rigor Rubric and Quality Index (Continued)

Domains	Checklist	Explanation	Responses	Source (Item Number)
	3.4. Were the costs measured and valued appropriately?	Average of three main components. 3.5.1, 3.5.2 and 3.5.3. Each sub question has equal weight in contributing to 3.5	0, 0.33, 0.67, or 1	Adams, 1992 (3,4,8); Evers, 2005 (8,9,14)
	3.4.1. Were the intervention cost measures used clearly described?	Direct measures (units and unit prices) or cost regression to identify marginal costs	0 or 1	Adams, 1992 (3,4); Evers, 2006 (8,9)
	3.4.2. Were monetized outcomes and intervention costs discounted when the costs were over a year? If not discounted was the reason explained?	1 if discounted. 1 if not discounted but reason was explained. 0 if no discounting and no explanation	0 or 1	Adams, 1992 (8); Evers, 2005 (14)
	3.4.3. Were costs adjusted to real values/ inflation?	1 if adjusted. 1 if not adjusted but reason was explained. 0 if no adjustment and no explanation	0 or 1	Evers, 2005 (9)
	3.5. Were control and treated (or cohorts) recruited from the same population?	1 if the answer is yes. 0 if groups from different population. 0 if there is no control group	0 or 1	Downs and Black, 1998 (21)
	3.6. Were control and treated (or cohorts) recruited from the same period?	1 if the answer is yes. 0 if the answer is no. 0 if the study period was not specified	0 or 1	Downs and Black, 1998 (22)
	3.6.1. What was the study period	Explicitly stated the years/months for the evaluated outcomes		
	3.7. Did the analyses adjust for different follow-up lengths in cohort or case-control studies? Or was the period between intervention and outcome the same for cases on controls?	1 if the follow-up was the same for all groups. 1 if different lengths were adjusted. 0 if the difference were ignored	0 or 1	Downs and Black, 1998 (17)
	3.8. Were attrition/ losses from follow-ups taken into account?	1 if it was mentioned and how it handled was explained. 1 if it was mentioned and was not handled but lost was too small to affect main findings. 0 if it was mentioned but was not handled. 0 if it was not mentioned.	0 or 1	Downs and Black, 1998 (9,26)

Table 2.A1. Methodological Rigor Rubric and Quality Index (Continued)

Domains	Checklist	Explanation	Responses	Source (Item Number)
4. External validity	4.1. Subjects participated in the study represents entire population from which they were recruited?	Validation that the sample was representative would include demonstrating that the distribution of the main confounding factors was the same in the study sample and the source population	0 or 1	Downs and Black, 1998 (11,12)
5. Power	5.1. Did study have sufficient power to detect a clinically important effect? (p-values)	Sample sizes have been calculated to detect a difference of x% and y%.	0 or 1	Downs and Black, 1998 (27)
Total points	Sum the values of reporting items (2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8), internal validity items (3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8), external validity item (4.1), and power item (5.1)		0-18	

Table 2.A2. Mean Effects on Recalculated ROI: Quality Index (Items Equally Weighted) and Article Characteristics (Excludes Outliers[1])

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Quality index: Items are equally weighted	-0.026						0.048
	(0.125)						(0.203)
Programs in a small company[2]		-0.168					0.319
		(1.166)					(1.743)
Only disease management rogram[3]			1.094				1.658
			(1.048)				(1.333)
Disease management and wellness programs[3]			0.551				0.507
			(0.958)				(1.272)
Only absenteeism costs[4]				-0.564			0.684
				(0.982)			(1.496)
Health care and absenteeism costs[4]				-0.884			-0.120
				(1.151)			(1.611)
Any productivity cost[4,5]				-1.760*			-0.980
				(0.926)			(1.287)
Publication year dummy (2011)[6]					-0.043		0.125
					(0.738)		(1.160)
Evaluation duration (at least 3 years)[7]						0.920	1.026
						(0.736)	(1.200)
Constant	1.136	0.843**	0.560	1.688**	0.848	0.245	-0.667
	(1.520)	(0.389)	(0.447)	(0.694)	(0.550)	(0.584)	(3.316)
Observations	27	27	27	27	27	27	27

Notes: 1. Extreme outliers of ROI which was below Q1-IQR*3 and above Q3+IQR*3 where Q1 is 25th percentile, Q3 is the 75th percentile, and IQR is the interquartile range
2. Small is 1 if the company has 500 and less employees, 0 otherwise
3. The base category is "Only wellness program"
4. The base category is only health care cost
5. This component includes combinations of costs that include productivity (presenteeism and else)
6 Year dummy is 1 if the study was published after 2010 (2011 and after)
7. Evaluation duration is 1 if the study period is at least 3 years

Table 2.A3. Mean Effects on Recalculated ROI: Reduced Quality Index and Quality Index Domains and Article Characteristics (Excludes Outliers[1])

	(1)	(2)	(3)	(4)
Reduce quality index	-0.038	-0.030		
	(0.142)	(0.228)		
Reporting (8 items)			-0.487	-0.851
			(0.662)	(0.861)
Internal validity (7 items)			0.094	0.168
			(0.237)	(0.303)
Programs in a small company[2]		-0.112		-0.553
		(1.760)		(1.817)
Only disease management program[3]		1.570		1.458
		(1.316)		(1.322)
Disease management and wellness programs[3]		0.440		0.098
		(1.262)		(1.310)
Only absenteeism costs[4]		0.562		0.703
		(1.499)		(1.506)
Health care and absenteeism costs[4]		-0.197		-0.103
		(1.612)		(1.616)
Any productivity cost[4,5]		-1.022		-1.262
		(1.290)		(1.314)
Publication year dummy (2011)[6]		0.296		0.439
		(1.153)		(1.163)
Evaluation duration (at least 3 years)[7]		0.940		0.794
		(1.184)		(1.194)
Constant	1.258	0.338	3.746	5.243
	(1.679)	(3.452)	(3.959)	(6.044)
Observations	27	27	27	27

Notes: 1. Extreme outliers of ROI which was below Q1-IQR*3 and above Q3+IQR*3 where Q1 is 25th percentile, Q3 is the 75th percentile, and IQR is the interquartile range
2. Small is 1 if the company has 500 and less employees, 0 otherwise
3. The base category is "Only wellness program"
4. The base category is only health care cost
5. This component includes combinations of costs that include productivity (presenteeism and else)
6 Year dummy is 1 if the study was published after 2010 (2011 and after)
7. Evaluation duration is 1 if the study period is at least 3 years

Table 2.A4. Mean Effects on Calculated ROI: Quality Index Some Subcomponents (Excludes outliers[1])

	(1)	(2)	(3)	(4)	(5)
Observational studies with control group[2]	-0.053			0.142	
	(0.976)			(1.129)	
Quasi-experimental[2]	1.265			1.150	
	(1.196)			(1.312)	
Randomized[2]	-0.828			-0.299	
	(1.000)			(1.194)	
Appropriate methods for outcome estimates		0.563		0.229	
		(0.829)		(0.938)	
Program cost measures are clearly described			-1.269	-0.893	
			(0.762)	(0.964)	
Conflict of interest					0.689
					(0.787)
Constant	0.900	0.678	1.200***	0.900	0.560
	(0.756)	(0.422)	(0.415)	(0.775)	(0.463)
Observations	27	27	27	27	26

Notes: 1. Extreme outliers of ROI which was below Q1-IQR*3 and above Q3+IQR*3 where Q1 is 25th percentile, Q3 is the 75th percentile, and IQR is the interquartile range
2. The base category is the observational studies without control group

CHAPTER III

WHY DO FIRMS IMPLEMENT WORKPLACE WELLNESS PROGRAMS? UNDERSTANDING THE REASONS BEYOND RETURNS ON INVESTMENT

Introduction

Avoidable behavior-related risk factors, such as physical inactivity, unhealthy eating habits, obesity, and high cholesterol have become major health issues that impose high costs on individuals, businesses, and governments. The total of direct costs (treatment costs) and indirect costs (lost income and productivity) of chronic illnesses reached $3.7 trillion in the U.S. in 2016, which was almost 20 percent of GDP (Waters and Graf, 2018). Government wants to improve population health to avoid high health care costs due to their share in government spending. Workplace wellness programs (WWPs) have been used to change health behavior to reduce risk factors and prevent chronic illnesses (WHO, 2005; Mattke et al., 2013; Edington et al., 2014). The U.S. government enacted the largest public investment in the U.S. for WWPs through the Patient Protection and Affordable Care Act of 2010 (ACA) to improve public health (Anderko et al., 2012; Mattke et al., 2013) and reduce health care costs. The ACA provides financial and technical support for the private sector in promoting health and evaluating WWPs. The U.S. National Institute for Occupational Health and Safety (NIOSH) launched the Total Worker Health (TWH) program in 2011. The TWH

establishes the persistent need for worksite research on WWP implementation and diffusion (NIOSH, 2012).

Investments in WWPs as public health policy tools require rigorous evaluations to provide evidence on the effectiveness of these programs. There are two main issues regarding the effectiveness of WWPs. The first is the assumption that ROI is the main organizational strategy for implementing WWPs. However, this narrow perspective excludes the alignment of organizational values and program outcomes to create healthier workplaces (Edington et al, 2016). The second is the critiques on aligning public and private perspectives on evaluated program outcomes. Encouraging policy-makers to invest in workplace wellness requires a clear understanding of the objectives and the outcomes of the programs from both public and private perspectives (Fielding, 1984).

Large companies with the objective of gaining positive returns to maximize profit have constituted the majority of economic evaluations. The ACA and NIOSH support WWP implementation and research in areas where the private market falls short. To receive and maintain government assistance, companies need to prove that WWPs are successful at promoting health and reducing health care costs. To support public health policy, businesses need to link a business' reasons of choosing WWPs to its program decision. We also can test the assumption that financial savings is the main reason for the implementation of WWPs. We also provide evidence on to what extent ROI explains WWP implementation.

Currently, there is no theoretical or empirical study on companies' decision-making processes on offering the most effective WWPs to accomplish their objectives

given the resources. This study is the first paper that develops an economic model to evaluate firms' behavior beyond profit maximization when choosing WWPs. The model allows for both profit maximizing objectives and objectives other than profit, such as utility from corporate citizenship. We maximize the overall objective function of a firm with respect to WWP choices and derive a system of demand equations for WWPs where we can look at the attributes of reasons for WWP implementations on each WWP choice. We estimate the system of equations using Well Workplace Checklist (WWC) data established by the Wellness Council of America (WELCOA). Using the estimates of WWP equations, we can link business' reasons of choosing WWPs to program decision. We also can test the assumption that financial savings drive the implementation of WWPs; provide evidence on to what extent ROI explains WWP implementation.

Background

WWPs need to be a part of organizational strategic plans to achieve organizational objectives. However, either the external wellness program providers or a companies' wellness teams tend to choose popular programs that exit in the wellness market. Objectives for WWP implementation depends on organizations' needs and characteristics (Mattke et al., 2013; Edington et al., 2014; Kaiser Family Foundation, 2017). For successful programs, WWP choices should reflect the needs and characteristics. A systematic review on WWPs in the U.S. finds that improving health and reducing health care costs are common evaluated outcomes (Mattke et al., 2013). This narrow focus does not provide any information about organization's corporate

strategies on social responsibility to create healthy workplace culture and environment (Kickbusch and Payne, 2003; Kirkland, 2014; Martinez-Lemos, 2015).

Organizational characteristics, such as industry, size, and profit structure (for-profit, nonprofit, and not-for-profit) influence the company's choices to achieve its objectives. Different industries are subject to different regulations and working conditions that determine the need and content of WWPs. The annual survey of the Kaiser Foundation in 2017 shows that service and manufacturing industries implement WWPs in higher proportions when compared to other industries. The size of the company affects the available resources, where large companies potentially have the advantage to generate resources for financial and logistic needs to implement and evaluate their WWPs (McCoy et al., 2014). The annual survey of the Kaiser Foundation in 2017 shows large companies, which has 200 and more employees, implemented WWPs have higher proportion compared to small companies. The systematic review of this dissertation shows that only 10 percent of included articles evaluated a WWP in a small company, which has less than 500 employees. Although the profit structure determines the objective for implementing WWPS, there is no research on how profit structure could affect the decision-making process.

In addition to aligning program objectives and program choice within organizations, aligning public and private value propositions is essential for public policy to achieve social welfare objectives of improving employee wellness. The alignment is relevant to government support in health care, such as the ACA in the US where the Act invested in private provision of WWPs to promote health and reduce risk factors.

Although research does not necessarily align either program choices with organizational objectives or with the value propositions of stakeholders, in practice we observe a shift where organizations view WWPs as a part of organizational strategy (Berry et al., 2010). To realize the full potential of WWPs, best practices to guide organizations were developed by academic and practitioner experts (O'Donnell et al., 1997; Goetzel et al., 1998; Chapman, 2004). Although different organizations have different reasons for wellness programs, establishing main pillars for best workplace wellness practices help employers to attain their program objectives.

Linkage to business objectives, leadership commitment, employee engagement, organizational and environmental support, and integrating relevant systems are some of the common areas identified by the following checklists: Well Workplace Checklist (WWC), Health Enhancement Research Organization (HERO), American Health Association (AHA), and National Workplace Health Promotion Survey (Chapman, 2004; Terry et al., 2008; Schill and Chosewood, 2013; Kent et al., 2016; O'Donnell, 2016).

The Well Workplace Checklist

We use data from WELCOA's WWC. The Wellness Council of America (WELCOA) was established as a national nonprofit organization in the mid-1980s. WELCOA provides resources for building high-performing healthy workplaces by enhancing employee well-being and improving organizational outcomes. In the 1990s, WELCOA initiated the Well Workplace Checklist (WWC), which is an online survey, to assess the quality of workplace wellness practices within organizations. The quality is assessed over seven benchmarks including capturing senior leadership (upper

management) support, creating wellness teams, data collection, operating plans, choosing proper interventions, supportive health promoting environment, and program evaluations. The WWC benchmarks provide insight into companies' wellness initiatives and lays out the characteristics for the best practices of workplace wellness. Beyond the availability of the data, the WWC is one of the most comprehensive efforts for the benchmarks.

The WWC includes questions about company characteristics and the benchmarks that WELCOA determined for best practices in workplace wellness. The demographics include questions about company size, industry, unionization, multiple sites, and multiple shifts. Other company characteristics include questions on companies' wellness programs. These questions include how long the wellness initiatives have been in place, how the WWPs are paid, participation rates, department which runs WWPs, annual budget for WWPs, and reason for WWP implementation.

There are seven benchmarks with a series of questions related to quality measures for senior leader (upper management) support, wellness teams, data collection, operating plans, programming, supportive environment, and evaluations. Each question within a benchmark was values between 0-100 where "0" is need improvement, "25" is good, "50" is very good, "75" is excellent, and"100" is superior. The focus of this paper is on the wellness programs. Thus, we only explain programming benchmark in the data section. The first benchmark, senior leader support, is to measure the strength of senior leadership to achieve successful WWP implementations. Questions evaluate CEO's communication, resource allocation, delegation, and health promotion practices. The second benchmark, wellness team, is to assess the quality of wellness team. A well-

functioning team is important to create and deliver effective programs. Questions measure wellness team's history and composition and operating methods. The third benchmark, data collection, evaluates organization's efforts for collecting data. Questions ask information on organizational data, employee data, and environmental data collections, and employee protection data. The fourth benchmark, operating plan, is to evaluate the communication as to what the program will accomplish. Questions ask whether the operating plan has the following: overall mission/vision statement, measurable objectives, and link wellness goals to strategic priorities and outcomes, timeline, responsibilities, budget, and evaluation of stated goals. The fifth benchmark, programming, is to assess organization's choice of health promotion interventions to its employees. The focus of this paper is the programming benchmark. We provide detailed information on questions for programming in the following subsection. The sixth benchmark, supportive environment, is to evaluate workplace environments that influence employees' decisions on healthier behaviors. Questions ask whether there is environmental support on physical activity like onsite facility, nutrition/weight management like healthy food options, occupational safety, stress management, prohibiting alcohol and drug use, restricting regulations for tobacco use, and providing wellness initiatives. The seventh benchmark, evaluations, is to assess wellness evaluation efforts. Questions ask whether the organization is tracking participation and satisfaction, assessing risk and biometric measures, measuring changes in the environment like working conditions, and analyzing economic outcomes.

Data

WELCOA provided data on organizations who voluntarily completed the WWC for the years 2008-2015. During these years, new organizations entered the sample, and some dropped out. The initial sample included 5,433 checklist entries as described in Weaver et al. (2016). Multiple entries of an organization within a year were excluded. We only kept the most recent entry within a year to avoid duplicates of an organization in a given year. Of the remaining, some entries that were completed for test purposes were excluded, as well as the entries that had missing demographic variables making it unclear as to whether it was a test or mock entry. International-based organizations were excluded from the sample because these organizations could be subject to legal regulations. In addition, there were only a very small number of international based organizations. Only an organizations' first entry to the checklist was included in this sample in order to avoid measurement issues that could arise due to experience in filing the survey. Companies that chose more than two reasons for WWP implementation were also excluded because: (1) the survey asks to choose top two reasons and (2) some companies picked up to 12 reasons out of 13 reasons which is not informative about the reason for implementation because they are not ranked. Thus, the final analysis sample includes 3695 companies (Figure 3.1).

Analysis Variables

There were 14 questions that asked which WWPs were implemented in the organization. The first two questions asked whether the organization had offered employees the opportunity to participate in a health risk appraisal (HRA) or a health

screening, respectively. Answers were categorized as “12 months,” “24 months,” “36 months,” “48 months,” and “not offered.” We generated binary variables of health risk appraisal and screening that took on value one if the answer was “12 months,” and zero otherwise. The remaining program-related questions, which are listed in the Background subsection, asked about the topic and formats of the organization’s WWPs over the last 12 months. Answers were categorized as “awareness,” “awareness and education,” “awareness, education, and behavioral change,” “awareness, education, behavioral change, and culture enhancement,” and “not offered.” We generated binary variables for each program that took on value one if the answer was “awareness,” “awareness and education,” “awareness, education, and behavioral change,” “awareness, education, behavioral change, or culture enhancement,” and zero otherwise.

There were thirteen reasons for WWP implementation that were included as response options on the checklist. The implementation reasons are listed as “improve employee health,” ”improve health of spouses and dependents,” “improve health of retirees,” “increase employee responsibility for managing personal health,” “contain health care costs,” “produce a return on investment,” “reduce unnecessary medical care utilization,” “increase work performance (reduce presenteeism),” “enhance productivity,” “reduce absenteeism,” “meet employees’ requests,” “attract and retain employees,” and “improve employee morale.” Organizations are asked to pick their top two reasons for implementing WWPs. A binary variable was generated for each reason separately that took on a value of one if the organization picked that reason. Companies were asked to select their top two reasons, in no particular ranked order. Thus, we excluded companies

that picked more than two reasons from the analysis to avoid measurement errors. To avoid small sample size issues in the analysis, we excluded health of retirees as a reason and grouped some of the reasons. Performance reason was grouped by including increase work performance (reduce presenteeism), enhance productivity, and absenteeism. Retention reason was grouped by including meet employees' requests, attract, and retain employees.

We used three company characteristics in our analysis: WWP experience in years, size, and industry. How long wellness programs have been in place is categorized as "just started," "1-3 years," "4-10 years," and "more than 10 years." We generated binary variables for three different experience levels. The first is *inexperienced (just started).* This captures companies that started offering WWPs less than a year ago at the year of checklist entry. The second is *1-3 years.* These companies offered WWPs 1 to 3 years before the year of checklist entry. The third is companies experienced *more than 3 years.* These companies offered WWPs at least four years before the year of checklist entry.

Company size is categorized as under 50, 50-100,101-500, 501-1000, 1001-4999, and 5000+. We generated a binary variable for small companies that took on a value of one if the company size was 500 or less employees, and zero otherwise. There were more than ten industries classified in the checklist. Industry includes manufacturing, services, communication, agricultural, construction, wholesale and retail, transportation, utilities, mining, and other (listed in the WWC). We generated two binary variables for manufacturing and service industries. We included only two main industries due to the small sample sizes in each industry.

The entry year of the checklist could affect the reason for WWP implementation and which program to implement because of trends and regulations at the company or government level. We generated a binary variable that took on a value of one if the entry year was 2011 or after, and zero otherwise. We chose 2011 to control for policy change due to the Affordable Care Act of 2010. We chose the year after the ACA to account for potential delays in implementing the policy in workplaces.

Descriptive Statistics

Table 3.1 shows the WWP experience frequencies. There were 1,102 inexperienced companies, 1,306 companies with 1-3 years of experience, and 1,287 companies with more than 3 years of experience. Table 3.2 presents descriptive statistics for reasons for WWP implementation and company characteristics. Improving employee health, which was selected by 50-60 percent of companies, and containing health care costs, which was selected by 60-70 percent of companies, were the top reasons for WWP implementation for all experience levels. Producing a ROI was selected by 3-5 percent of companies depending on the WWP experience level. Around half of the companies in the checklist were in the service industry. Seventy-one percent of just started companies were small, whereas around 40-50 percent of companies with some WWP experience were small companies. Around half of the sample entered the checklist after 2010.

Figures 3.2-3.4 present the shifts in the WWP implementation reasons over time based on WWP experience levels. Figure 3.2 shows that containing health care costs as reason for WWP implementation has been decreasing in frequency for inexperienced and 1-3-year experienced companies. Change in frequencies is higher in the inexperienced

sample compared to experienced companies. Overall, ROI and health care utilization are less popular reasons for WWP implementation for any experience level. However, we observe a slight decrease in frequencies for inexperienced companies after 2013 and an increase in frequencies for experienced companies.

Figure 3.3 shows that improving employee health is the most popular health reason for WWP implementation. All the health reasons remained consistent over the years other than slight changes in frequencies of employee health reason. Figure 3.4 shows the trends for performance, retention, and morale reasons. Performance as a reason slightly increased in frequencies after 2013 for inexperienced and 1-3-year experienced companies. Employee morale and retention as reasons for WWP implementation increased in frequencies after 2012 only for inexperienced companies. Performance, retention, and morale reasons remained consistent over years for companies with more than 3-year experience.

Table 3.3 depicts the descriptive statistics for implemented WWPs. Physical activity and nutrition/weight management programs were the most preferred WWP programs. Over 80 percent of companies in full sample, over 50 percent of companies in the inexperienced sample, and over 90 percent of companies in the more experienced sample selected one of these two programs. Stress management and safety/health protection were the other highly preferred programs for the full sample. Overall, means for WWPs for the inexperienced sample were lower compared to the WWP experienced samples.

Figures 3.5-3.8 present the patterns for the most preferred wellness programs over the reasons for WWP implementation. We picked the most preferred health and cost reasons as well as employee morale to provide more insight for the descriptive information. Physical activity and nutrition/weight management programs were preferred by most of the companies regardless of the reason and experience level. If the reason for implementation was to contain health care costs, most of the inexperienced companies preferred offering smoking cessation program and most of the experienced companies preferred offering HRA and biometric screenings. If the reason for implementation is improving employee health, most of the inexperienced companies preferred offering smoking cessation program; most of the experienced companies preferred offering stress management programs. If the reason for implementation was to improve employee morale, most of the inexperienced and 1-3-year experienced companies preferred offering stress management and smoking cessation programs; and most of the more than 4-year experienced companies preferred offering stress management programs and biometric screening.

The descriptive statistics show that some programs were preferred regardless of the objectives of WWPs. The WWP choices are likely to depend on the perceptions of employers and WWP vendors such as insurance companies. This subsection provides a quick insight on what we expect from the data. In the next section, we analyze the association between the program choice and reasons for implementing WWPs using an econometric model by modeling the WWP demand to better understand the real-world practices.

Methods

Theoretical Model

One of the shortcomings of the economic evaluation literature of WWPs is the lack of any connection between implementing WWPs and the organizations' objectives. The evaluations have been executed from employer's perspective. Hence, profit-maximizing motives, cost savings and ROI, have been assumed as the primary reasons for WWP implementation. The theoretical model in this paper shifts the focus in two ways. First, it shifts from profit maximizing motives to an overall objective function of an organization that has profit and other components that are not necessarily profit related. Second, it shifts from returns or cost savings as the primary reason to other potential reasons, such as employee morale, retention, and employee request.

This section sets up the optimization problem of a representative firm that chooses WWPs to achieve organizational objectives. For simplicity, we assumed that there are three reasons for WWP implementation (*R)* and two WWPs ($x \epsilon \{ x_1, x_2 \}$) that the firm can choose to implement to achieve its objectives. Reasons for WWP implementation are savings from health care and productivity- related costs (*S*), improving employee health (*H*), and improving employee morale (*M*). x_1 and x_2 represent the two unspecified WWPs that a firm can chose between.

$$R \in \{S, H, M\}$$
$$x \in \{x_1, x_2\}$$

The program choice is assumed to attain the reason for program implementation. Thus, each reason for WWP implementation is a function of WWP choices:

$$S = s(x) \quad (3.1)$$
$$H = h(x)$$
$$M = m(x)$$

The decision-making process for an organization depends on its objective function, which is the utility function (u) of a firm. We use a random utility model to establish a functional relationship between program options, implementation reasons, and organizations' characteristics. The objective function has a systematic utility (V), and a random residual (e). The systematic utility is the perceived mean utility and the random residual is the unknown deviation from the perceived utility (Train, 1986; Cascetta, 2009). The systematic utility is a function of profit (π) and other utility components (ϕ) that are not profit related.

$$u = V(\pi, \phi) + e \quad (3.2)$$

The profit component, which is standard from economic theory, is the difference between total revenue and total cost. Total revenue is calculated as price *(p)* times production *(Y).* Total cost is determined with cost function *(C).*

$$\pi = pY - C \quad (3.3)$$

Production is the function of labor input (L), because WWPs impact only labor input in the production process. In this model, labor is a function of employee health where better health increases labor productivity.

$$Y = f(L) \tag{3.4}$$

$$L = L(H) \tag{3.5}$$

Cost function includes labor cost, workers' compensation claims, and program cost. Wages (w), health care benefits (b), and sick days (d) determine the labor cost. The labor cost is the sum of health care benefits (bwL), cost of lost days (absenteeism) (dwL), and cost of presenteeism (τwL). If there is no productivity loss due to presenteeism, then τ =1 and the cost is wL. If there is loss due to presenteeism, then τ >1 and the cost is τwL. Worker's compensation claims ($comp$) are the cost of work-related injuries. If any program is chosen given the reason for program implementation, represented by the indicator function ($I(x|R)$), then the cost function includes program costs (pc_x).

$$C = (\tau(x) + b(x) + d(x))wL(h(x)) + \theta_1 comp(x) + \theta_2 pc_x\{I(x \mid R)\} \tag{3.6}$$

The utility components that are not profit related (ϕ) create a corporate citizenship value for organizations to achieve firm's goals (Peredo and McLean, 2006). Employee health and employee morale both affect this component, which is the additional piece to the classical firm theory of profit maximizing objectives. Including

this component helps us to identify different motivations of for-profit, not-for profit, and nonprofit companies.

$$\phi = \phi(H, M) \tag{3.7}$$

Savings from health care benefit payments, presenteeism, absenteeism, or injuries affect firm's objective function through profit. Improving employee morale affects objective function through the utility component other than profit. In addition, improving employee health affects the objective function through both the profit and the other utility components. Enhanced employee performance is expected through improved labor productivity, presenteeism, and absenteeism.

The utility component that is not profit (ϕ) has a direct impact on the objective function. There are two reasons for implementation that could affect this component: improving employee morale and health. Employee health can affect the model in different ways depending on the employers' motives. Employee health enters the firm's utility through multiple mechanisms. The first mechanism is the impact of employee health on productivity shown in equations (3.4) and (3.5). The second mechanism is on cost via absenteeism, health care benefits, and workers' compensation claims shown in equation (3.6). The third mechanism is the impact of employee health on ϕ shown in equation (3.7)

Suppose that the objective function has the following linear form with parameters α and β that are organization's weight on profit and other utility components:

$$u = \alpha\pi + \beta\phi + e \tag{3.8}$$

An organization maximizes its utility with respect to labor input and WWP choices:

$$\max_{x\in\{x_1,x_2\},L>0}\{\alpha\pi + \beta\phi + e\} \tag{3.9}$$

We plug equations (3.1) and (3.3)–(3.6) into equation (3.8) and maximize it with respect to labor input and program choices. One of the primary objectives of the model is to analyze how reasons for WWP implementation links to program choice given the company characteristics using economic theory. Implicit function theorem allows us to obtain implied demand functions for WWP options that are functions of reasons and company characteristics. This paper does not solve the theoretical model for the choice variables explicitly. The first order conditions to the optimization problem and other relevant information to the model are provided in the Appendix 3.A. Estimation strategy for program choices is explained in the empirical model section.

The model has important assumptions for simplification. Each reason for implementation could have an indirect influence on the other reason. Improving employee morale could influence employee health and presenteeism. Improving employee health could influence health care utilization, presenteeism, absenteeism, and morale. The first assumption is that the reasons are independent from each other due to prioritizing primary reason for WWP implementation. In addition, once an organization chooses the reason, it implements the relevant program to achieve its organizational

objectives. The second assumption is that organizations offer the most effective WWPs based on their reason for program implementation to maximize their objective function.

Empirical Model

In practice, organizations might not achieve their stated objectives. As noted in the theoretical model section, the model assumes that organizations offer the most effective WWPs based on their reason for program implementation to maximize their objective function. However, the real-world practices have shown that this assumption fails and, in most cases, there is no effective association between reason for WWP implementation and the offered program. This issue reduces the reliability of WWPs as policy tools. The empirical model examines employers' perception on the implementation reasons and WWPs. This set up allows us to explore the disconnection between the business practices and evaluation literature.

The model estimates whether implementation reasons and organizational characteristics, such as WWP experience, industry, organization size and year of entry significantly influences choosing certain WWPs. The first step is to look to at the estimations for each program:

$$WWP_{ij}^{*} = X_{ij}\beta_{ij} + \varepsilon_{ij} \tag{3.10}$$

$$WWP_{ij} = \begin{cases} 1 & if \quad WWP_{ij}^{*} > 0 \\ 1 & if \quad WWP_{ij}^{*} \leq 0 \end{cases} \tag{3.11}$$

$$X_{ij} = [reason_{ij}, sector_{i}, small_{i}, d2011_{i}] \tag{3.12}$$

We first estimated the single choice probabilities of an organization choosing each program at a given time and the measured characteristics of observations using the probit model. Each program choice, $j \in \{1, 2, \ldots, 15\}$ is a function of X_{ij} that includes adoption reasons, organizational characteristics (company size, and industry) and dummy for checklist entry year of 2011 and after. Each error term, ε_{ij}, is distributed as standard normal with zero mean and variance one.

Each WWP option of *j* for organization *i* (WWP_{ij}) is a binary variable that takes on a value one if an organization chooses to implement it, and zero otherwise. Organizations can choose from the following WWP options: not to offer any program (no WWP), or to offer programs including (1) health risk appraisals, (2) health screenings, (3) physical activity, (4) smoking cessation, (5) nutrition and weight management, (6) alcohol consumption, (7) stress management, (8) medical self-care, (9) work and family, (10) personal financial management, (11) health and safety protection, (12) ergonomics, (13) mental health and depression, or (14) disease management.

Reason for WWP implementation ($reason_{ij}$) is a binary variable that takes on a value one if organization chooses it, and zero otherwise. Organizations choose their top two reasons for adoption, which were not ranked. The reasons were listed as (1) improve employee health, (2) improve health of spouses and dependents, (3) increase employee responsibility for managing personal health, (4) contain health care costs, (5) produce a return on investment, (6) reduce unnecessary medical care utilization, (7) improve work performance (reduce presenteeism, enhance productivity, or reduce absenteeism), (9)

retention (meet employees' requests or attract and retain employees), and (10) improve employee morale. The grouping accounted for the relevancy among the grouped reasons. Furthermore, improving retiree health was excluded from the repressors because of small sample size issue.

Sector ($sector_i$) is another variable that potentially affects the WWP options due to the relevant regulations and responsibilities. There are two binary variables that were used in the estimations: manufacturing and services. The other categories were considered the base category. We chose these two categories for sectors because they had the highest two subsample sizes. Company size ($small_i$), whether it is small or not impacts the resourced available and needs for WWP.

The year variable in the checklist is the first year that companies filled out the checklist. We used a dummy variable that is one if the checklist year was 2011 and after. We picked dummy for 2011 ($d2011_i$) to account for the ACA effect on the reason for WWP implementation and the WWP choice.

Univariate probit estimations considers all the program choices independent from each other. In the data, we observe that companies chose more than one program to offer. Equation-by-equation estimations impose zero restriction on coefficients of other equations. We also estimate the choice probabilities using the multivariate probit (MP) model to estimate joint probabilities of choosing more than one program. Acknowledging the potential covariance between program choices improves efficiency of the coefficient estimates (Zellner, 1962; Zellner and Huang, 1962). Error terms, ε_{ij} , are distributed as

multivariate normal with zero means and variance matrix where the diagonal (variance) takes on a value of one and off diagonal (covariance) takes on a value of ρ_{kj} for $j \neq k$.

Stata Statistical Software (Release 14) is used to estimate the MP model. The model estimated the following log likelihood function with twenty draws using a sample of N-independent observations (Cappellari and Jenkins, 2003):

$$\log L = \sum_{i=1}^{N} \log \Phi_{14}(\mu_i, \Omega) \tag{3.13}$$

$$\mu_i = (k_{i1}X_{i1}\beta_{i1}, ..., k_{i14}X_{i14}\beta_{i14})$$

$$k_{ij} = 2WWP_{ij} - 1 \qquad \forall i$$

$$\Omega = \begin{bmatrix} 1 & \cdots & k_{i1}k_{i14}\rho_{i1,14} \\ \vdots & \ddots & \vdots \\ k_{i1}k_{i14}\rho_{i1,14} & \cdots & 1 \end{bmatrix}$$

where $\Phi_{14}(.)$ is the multivariate normal cumulative distribution for 14 dependent variables (reasons for adoption) with mean μ_i and variance-covariance matrix Ω. We found that the estimated correlations between error terms are statistically significant and demonstrate the strong relation of these equations. Thus, we provided estimates of the MP model that allows firms to pick multiple programs and eliminates cross-equation restrictions.

As shown in the Data section, we observe shifts in reasons for WWP implementation that differed based on WWP experience. Although industry and size

influence the decision-making process for WWP implementation, we did not observe shifts in reasons for WWP implementation trends based on industry and size. Thus, regressions were run separately for experience levels and each regression we control for industry and size. Also, we cannot add interaction terms of reasons for WWP implementation and experience levels due to the number of right hand side variables that the multivariate probit can handle.

Results

Tables 3.4-3.6 presents the results for the empirical model. Table 3.4 presents the results for inexperienced companies. If the reason for WWP adoption was to improve health of dependents and spouses or to improve employee morale, companies were more likely to implement personal financial management programs. If the reason was to increase responsibility in managing personal health, then companies were more likely to offer health screenings or implement work family or personal financial management programs. Lastly, if the reason was containing health care costs, companies were more likely to offer health screenings. Companies in the service industry were less likely to adopt health/safety protection programs, while companies in manufacturing industry were more likely to adopt smoking cessation, health/safety protection, or ergonomics programs. Small companies were always less likely to adopt any type of wellness program.

Table 3.5 presents the results for companies with 1-3 years of experience companies. If the reason for WWP implementation was improving employee health, improving health of dependents and spouses, containing health care costs, producing

ROI, or improving work performance, companies were more likely to implement smoking cessation programs and less likely to implement ergonomics programs. If the reason was increasing responsibility for managing personal health, reducing medical care utilization, or improving employee morale, companies were less likely to implement ergonomics programs. If the reason was retention and employee request, companies were more likely to implement smoking cessation programs. Companies in the service industry were more likely to implement ergonomics programs, and less likely to implement physical activity, nutrition/weight management, and health/safety protection programs. In comparison, companies in the manufacturing industry were more likely to implement nutrition/weight management, health/safety protection, or ergonomics programs, and less likely to implement physical activity, stress management, and work family programs. Again, Small companies were always less likely to implement any wellness program.

Table 3.6 shows the results for companies with more than 3 years of experience companies. If the reason for WWP implementation was improving employee health, increasing responsibility for managing personal health, containing health care costs, or retention, companies were more likely to implement physical activity programs, and less likely to implement stress management programs. If the reason was to improve health of dependents and spouses, then companies were more likely to offer HRA and biometric screening and implement physical activity, smoking cessation, responsible alcohol use, ergonomics, and disease management programs. If the reason was to produce ROI, companies were more likely to offer HRA and implement physical activity programs, and less likely to implement stress management programs. If the reason was to reduce health

care utilization, companies implemented less stress management programs. If the reason was to improve employee morale, companies were more likely to implement physical activity programs. Companies in service industry were more likely to implement stress management and mental health/depression programs, and less likely to implement disease management programs. However, companies in manufacturing industry were more likely to implement smoking cessation or health/safety protection programs. As before, small companies were always less likely to implement any wellness program.

Significance of coefficients do not change much between univariate and multivariate regressions for inexperienced sample, and for size, industry, and year dummy variables in all subsamples. However, allowing correlation between equations improves efficiency of estimates for experienced samples. In 1-3-year experienced sample, we observe trends in ergonomics with multivariate regressions but not with univariate regressions. Similarly, in more than 3-year experienced sample we observe trends in stress management with multivariate regressions but not with univariate regressions. (Appendix 3.B1-3.B3). All pairwise correlations between equations in multivariate regressions are significant.

The results show that companies that have more than 3 years of experience prefer more physical activity programs and less stress management programs with most of the reason for WWP implementation. In addition, companies with 1-3-years of experience prefer more smoking cessation and less ergonomics programs with most of the reasons. The inexperienced sample does not have a clear pattern with choice of WWPs. Company characteristics show similar patterns across all experience levels. Small companies are

always less likely to implement WWPs. Companies in the service industry are less likely to implement health/safety protection programs. Companies in the manufacturing industry are more likely to implement health/safety protection and ergonomics programs. Nutrition/weight management programs, HRAs, and biometric screenings seemed to be the most popular programs in the descriptive statistics (Figures 3.5-3.8). Nevertheless, we no longer find that pattern when we control for all the reasons for implementation with company characteristics and allow for correlation between program choices in the regressions.

Our findings provide significant evidence that there are clear trends for WWP choices based on the experience and company characteristics, but that there is a disconnect between reasons for program implementation and program choice, especially with less experienced companies. Something important to note is that companies with 1-3 years of WWP experience implemented certain programs regardless of the program objective. Moreover, ROI is not necessarily the main value of WWPs to companies and the current state of the literature does not align the evaluation metric with true company values.

Discussion

Economic evaluations on WWPs have mainly focused on ROI analysis. These analyses demonstrate that WWPs could help companies receive financial savings through reduced health care costs and improved labor productivity. Most studies that evaluated WWPs did not provide information on the reasons for why companies would implement these programs other than ROI or cost savings. Although the success criteria for WWPs

need to be linked to businesses' strategies to provide valid and reliable evaluations, there are no theoretical or empirical studies on companies' decisions to adopt WWPs.

The findings of this paper show that inexperienced companies did not have a strategy on how to utilize WWPs to achieve their objectives. As the experience for WWP implementation grows over 1-3 years, companies were more likely to choose smoking cessations programs. Employers chose smoking cessation programs to achieve health improvements, cost savings, or other objectives because health insurance providers offer those programs. These companies were less likely to implement ergonomics programs. Most insurance companies offer smoking cessation programs for free (no cost on employers). Preferences of insurance companies might be the main reason of these trends. In addition, companies might already have satisfied the regulations relevant to ergonomics and there was no need for ergonomics programs. As the experience for WWP implementation increases to more than 3 years, companies were more likely to implement physical activity programs. Employers chose physical activity programs also to achieve certain health improvements, cost savings, or other objectives. These companies were less likely to choose stress management programs. As experience with WWPs increases, companies might prefer more complex programs. That might be why we observe a shift from smoking cessation (for 1-3 years of experience) to physical activity programs (for more than 3 years of experience). Moreover, companies with the most experience might want tangible results which are relatively easier to evaluate. That might be the reason for offering physical activity programs but not stress management programs.

Companies often work with the insurance providers for WWPs implementation. Looking at these patterns, it seems that companies may follow the WWP trends that health insurance companies and wellness program vendors provide. This is suggested by the lack of association between objectives for WWP implementation and program choice. When health insurance providers offer WWPs, they usually present descriptive information on participation and certain health information. Suppose the objective for implementing WWPs is improving employee morale and physical activity is offered to achieve the employers' objective. Then, employers should have data collection that suggests that their employees want physical activity programs and evaluate the effect of these programs on morale. Outside vendors or insurance companies may only measure what is important to them, not the employers. Therefore, even if there is strong public policy support, desired outcomes may not be achieved due to missing connections. Research that examines the relationship between value propositions of and programs provided by vendors or insurers may help to identify misalignments in value propositions and help companies choose programs that may be more likely to meet their objectives.

There are four limitations relevant to this study. First, the checklist is self-reported where the data is subject to measurement errors. Second, the position of the respondent who completed the checklist on behalf of the organization could influence responses, particularly for value propositions of WWPs. For instance, managers and employee-wellness team might have different reasons for WWP implementation. Although future research can analyze these differences in perceptions by looking at variations in value propositions across different levels of employees and employers in organizations, it is not

the focus if this paper. Third, there may be recall bias in responses, especially for checklist questions on WWPs which asked whether the organization implemented that WWP in the last 12 months. Last, the checklist asked for the top two reasons for WWP implementation in no particular ranked order. Although this question gives more detail than asking to choose all that apply, we still cannot identify the top reason that could affect the regression results in the model.

Although there are limitations, this research offers a new perspective on how to improve successful adoption and assessment of wellness programs in workplaces with different company characteristic. The economic evaluations have presented ROI as the targeted outcome of WWPs. Although information on returns show the use of resources from a business perspective, the success of a WWP to achieve certain objectives cannot be tied to this narrow perspective. This is the first and so far only study that evaluates value propositions from the economics perspective and models the firms' behavior when choosing WWPs using economic theory.

Figures

Figure 3.1. Sampling for WELCOA Analysis

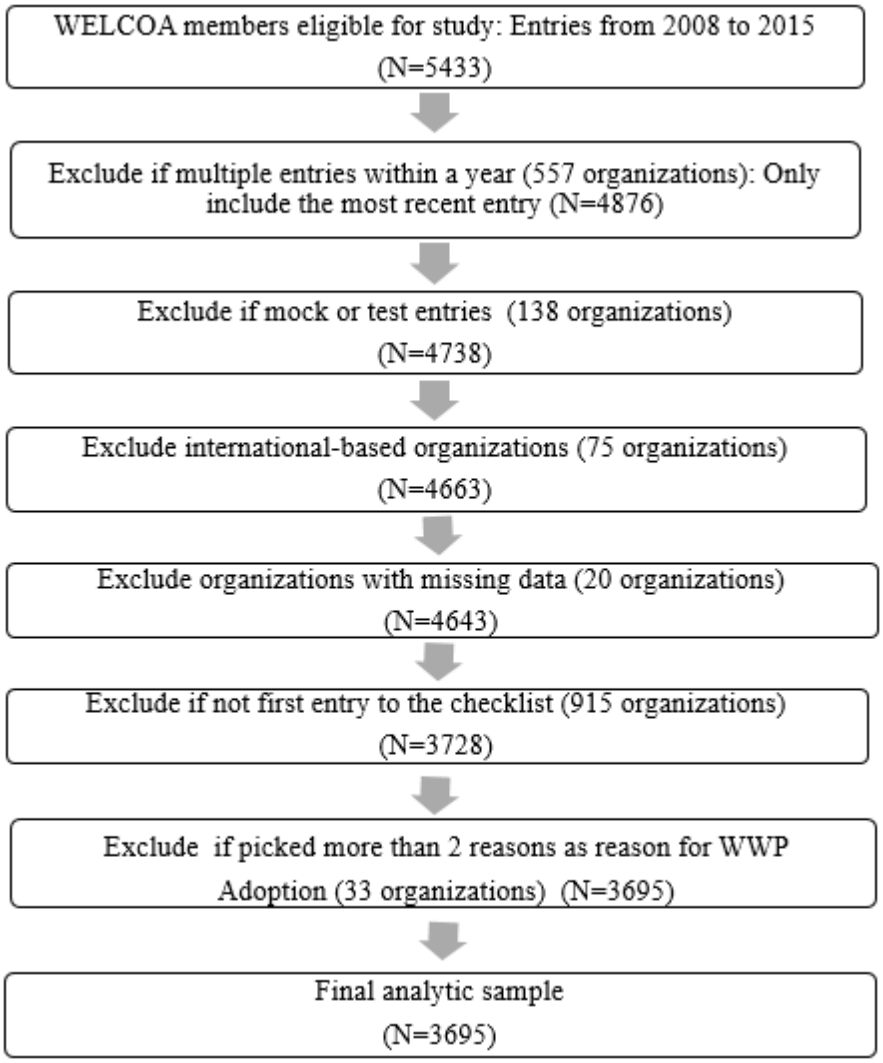


Figure 3.2. Sample Percentages of Cost Reasons for Implementing WWPs by Experience

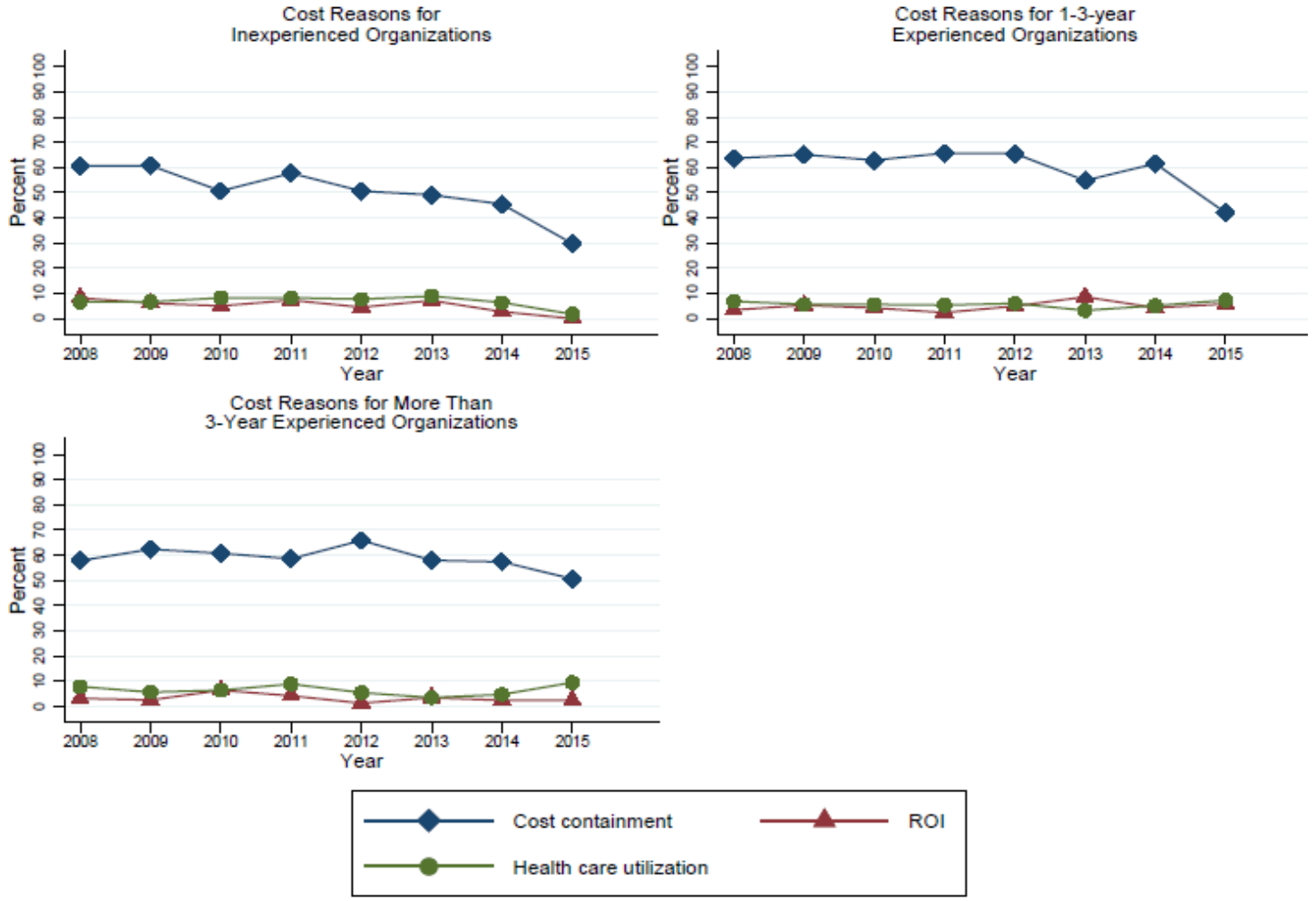


Figure 3.3. Sample Percentages of Health Reasons for Implementing WWPs by Experience

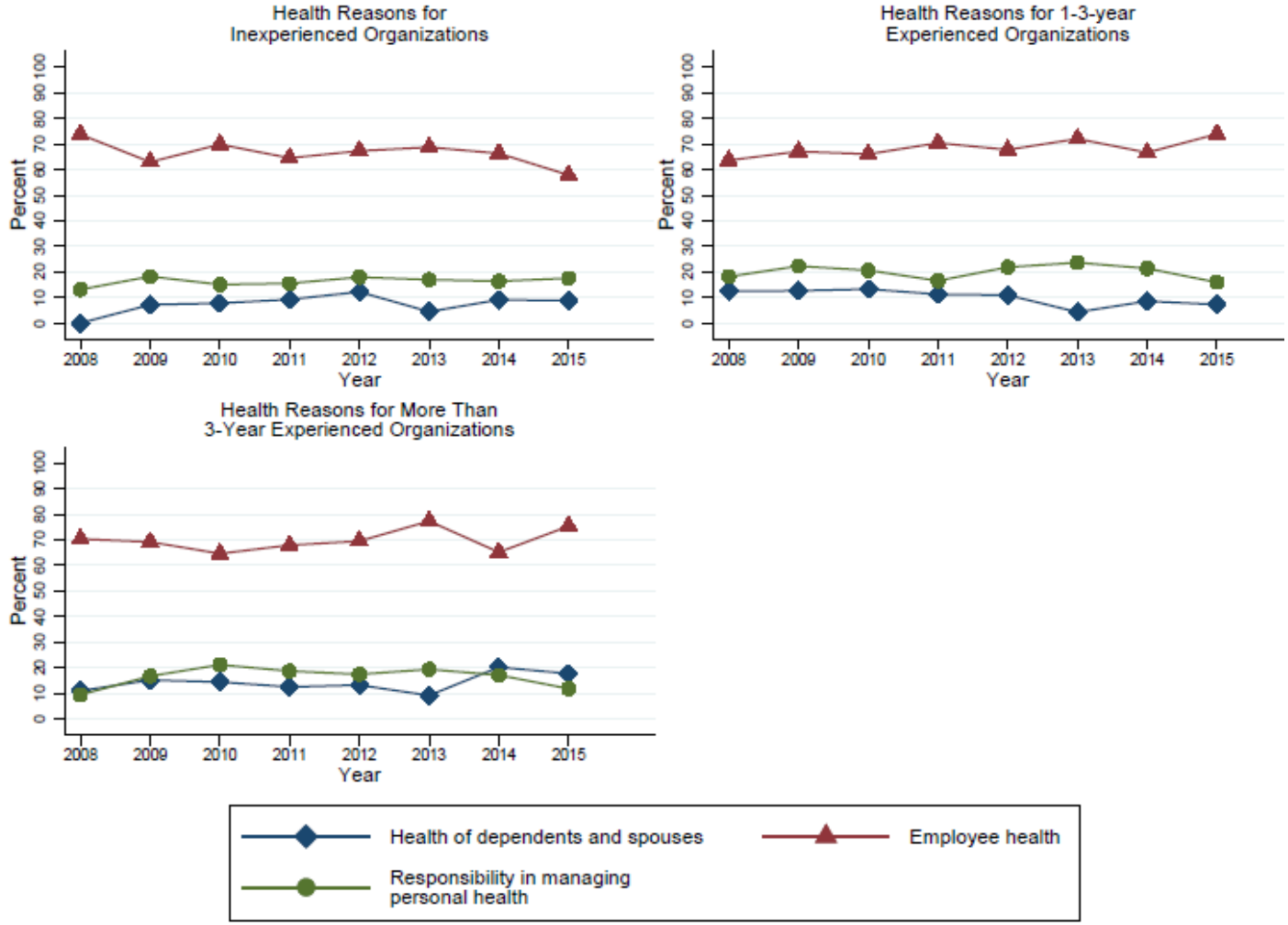


Figure 3.4. Sample Percentages of Reasons Other than Cost and Health for Implementing WWPs by Experience

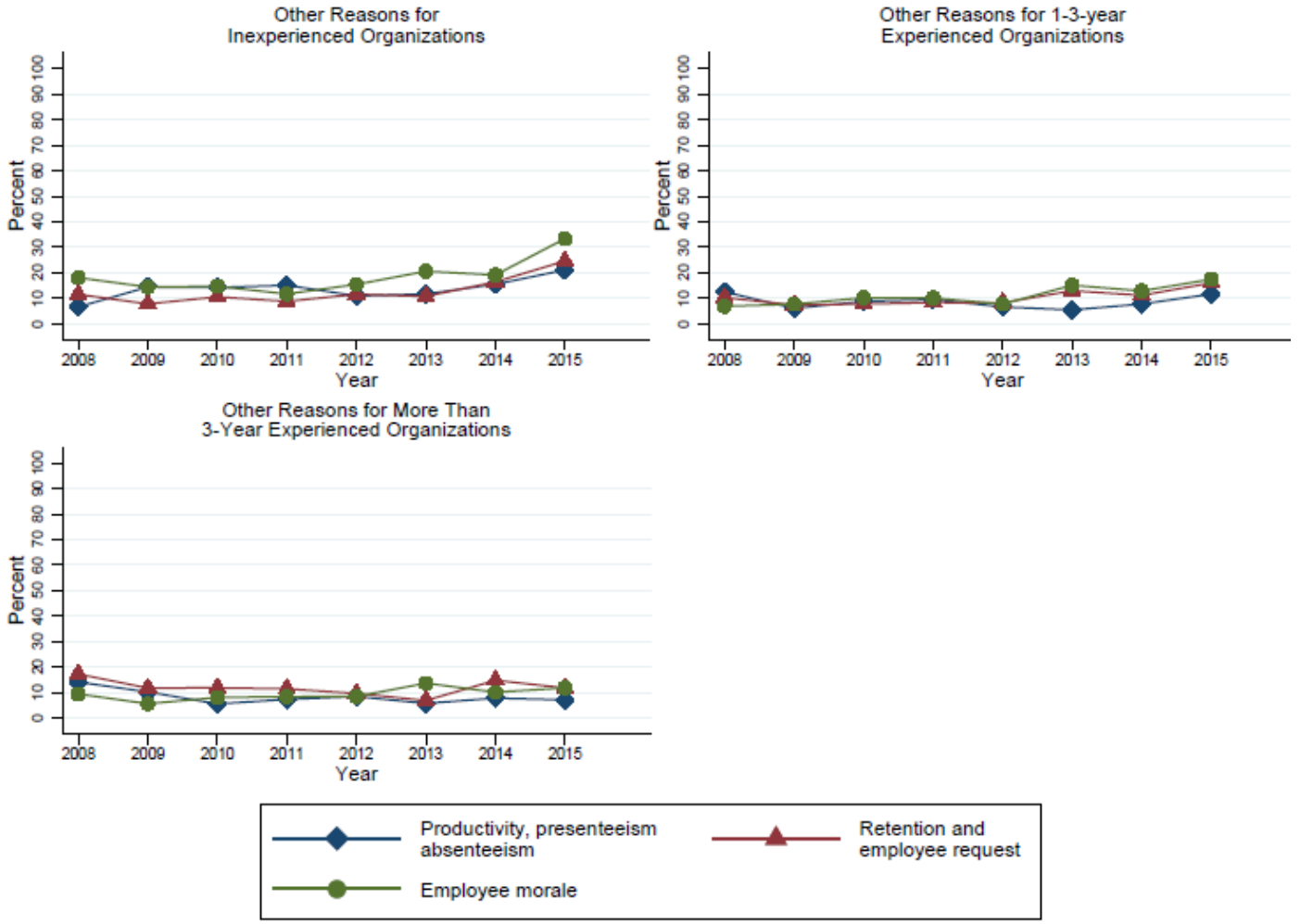


Figure 3.5. Frequenies for the HRA, Health Screening, and Most Offered Wellness Programs by Experience: Containing Health Care Costs

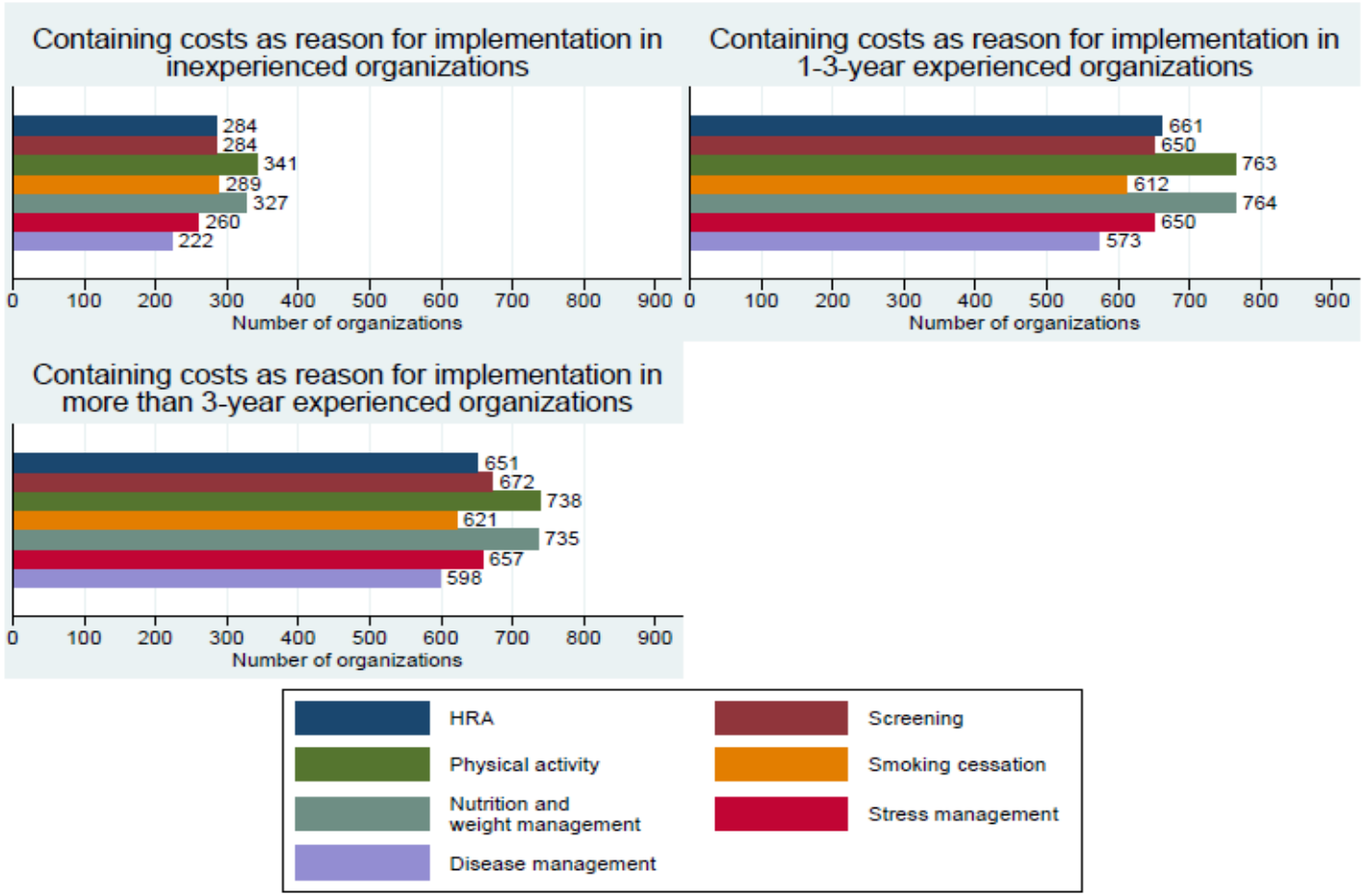


Figure 3.6. Frequenies for the HRA, Health Screening, and Most Offered Wellness Programs by Experience: Improving Employee Health

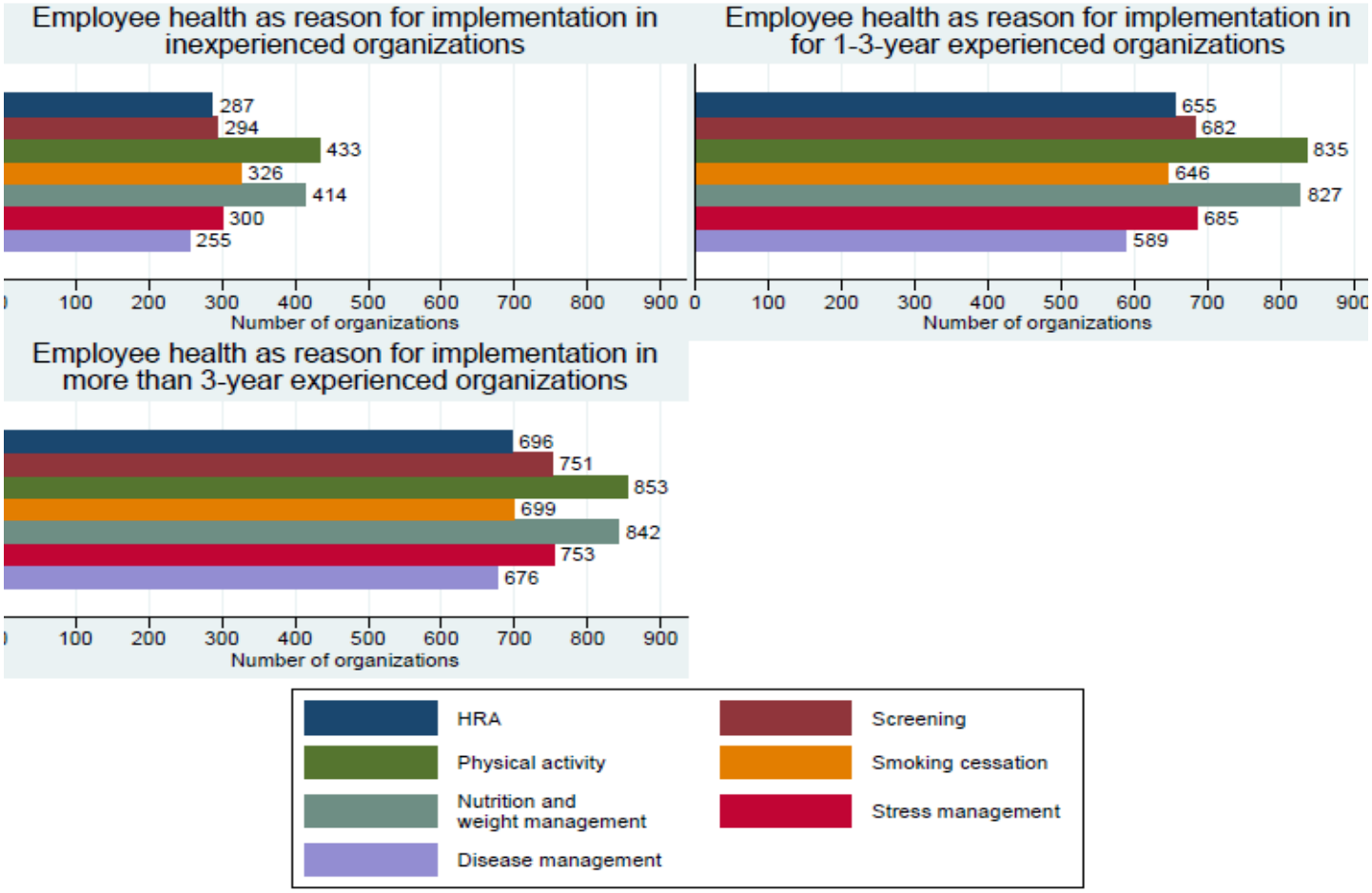


Figure 3.7. Frequenies for the HRA, Health Screening, and Most Offered Wellness Programs by Experience: Improving Employee Performace

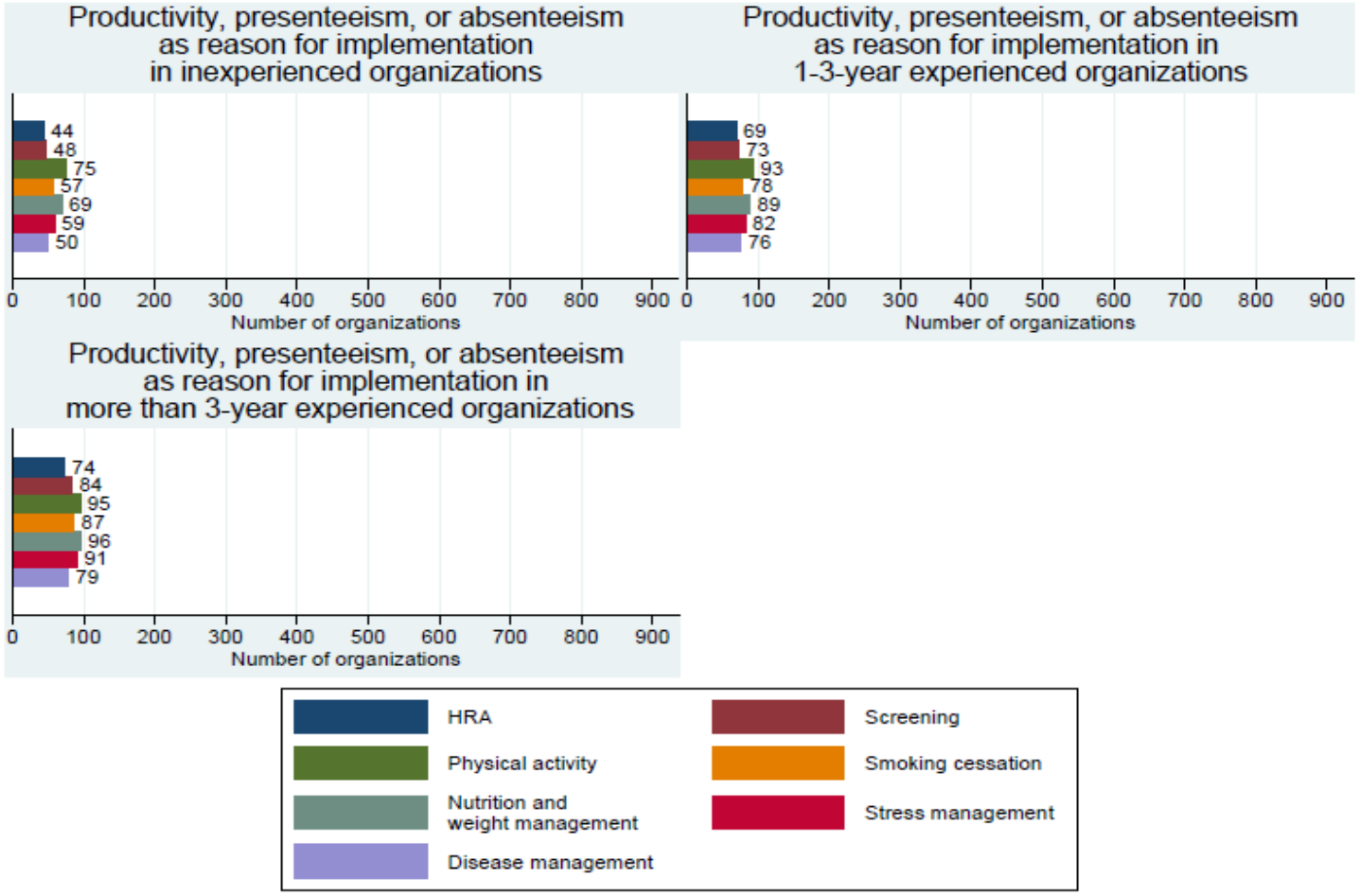
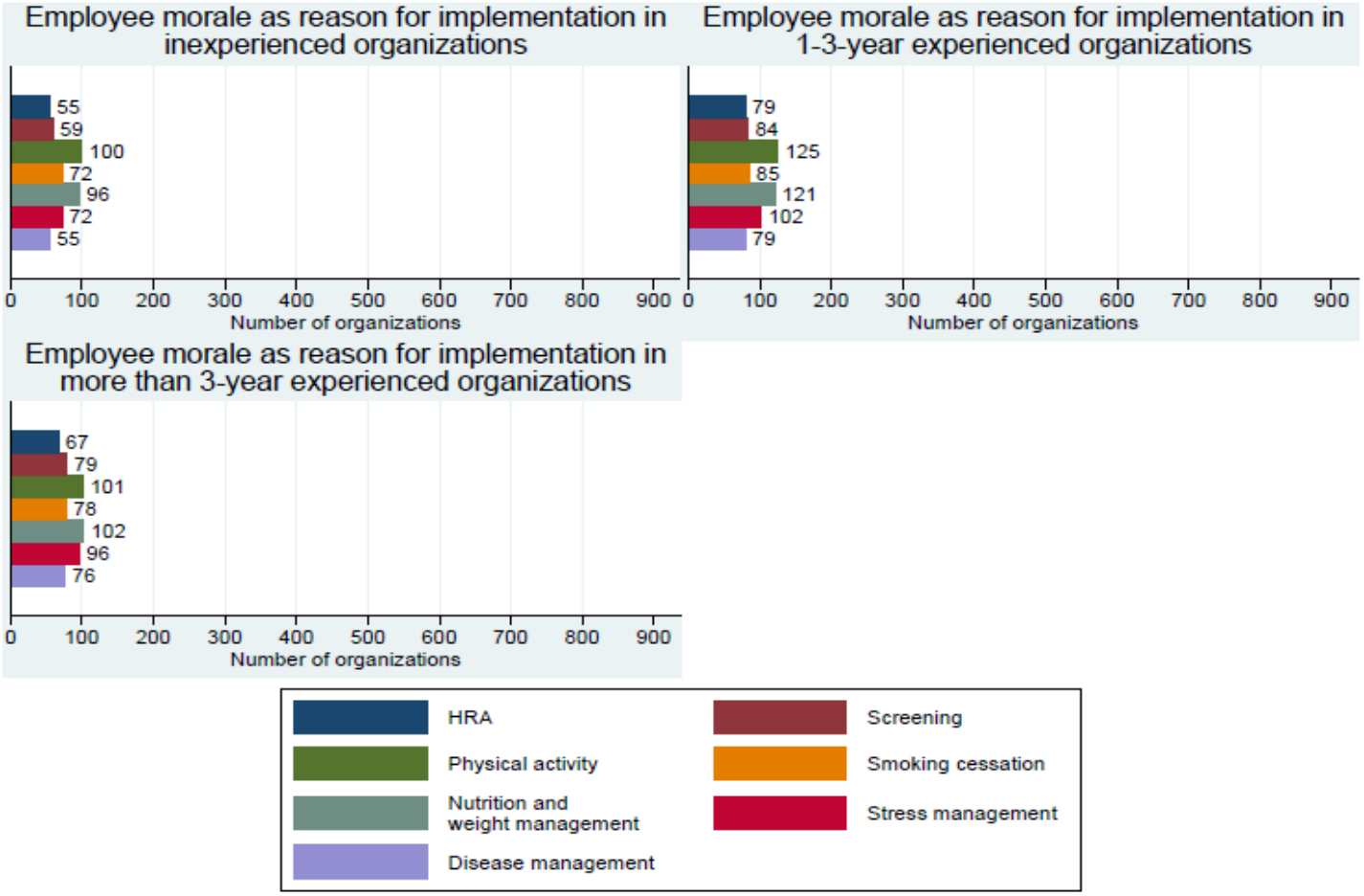


Figure 3.8. Frequenies for the HRA, Health Screening, and Most Offered Wellness Programs by Experience: Improving Employee Morale



Tables

Table 3.1. Frequencies for How Long Wellness Programs Had Been in Place

Experience in implementing wellness program	Frequency	Percentage
Program has just been in place	1102	29.82
Program has been in place for 1-3 years	1306	35.35
Program has been in place for more than 3 years	1287	34.83

Table 3.2. Descriptive Statistics for Explanatory Variables

	(1)	(2)	(3)	(4)
	Full sample	**Inexperienced**	**1-3-year experience**	**More than 3-year experience**
Reasons for WWP Adoption				
Improve employee health	0.678	0.665	0.678	0.688
	(0.467)	(0.472)	(0.467)	(0.464)
Improve the health of employees' spouses and dependents	0.113	0.080	0.111	0.144
	(0.317)	(0.271)	(0.314)	(0.351)
Increase employee responsibility for managing personal health	0.182	0.164	0.205	0.174
	(0.386)	(0.371)	(0.404)	(0.379)
Contain health care costs	0.586	0.525	0.624	0.601
	(0.493)	(0.500)	(0.485)	(0.490)
Produce ROI	0.044	0.054	0.047	0.033
	(0.206)	(0.227)	(0.211)	(0.180)
Reduce unnecessary medical care utilization	0.064	0.074	0.056	0.063
	(0.244)	(0.261)	(0.230)	(0.243)
Improve performance or productivity	0.075	0.101	0.062	0.067
	(0.264)	(0.301)	(0.241)	(0.250)
Reduce absenteeism	0.026	0.044	0.018	0.018
	(0.159)	(0.206)	(0.134)	(0.133)
Employee request or retention	0.106	0.113	0.090	0.117
	(0.308)	(0.316)	(0.287)	(0.321)
Improve employee morale	0.113	0.163	0.100	0.084
	(0.317)	(0.370)	(0.300)	(0.277)
Company Characteristics				
Industry: Services	0.499	0.545	0.499	0.458
	(0.500)	(0.498)	(0.500)	(0.498)
Industry: Manufacturing	0.154	0.145	0.157	0.159
	(0.361)	(0.352)	(0.364)	(0.365)
Industry: All other industries	0.347	0.309	0.344	0.383
	(0.476)	(0.462)	(0.475)	(0.486)
Size: Small	0.572	0.710	0.565	0.462
	(0.495)	(0.454)	(0.496)	(0.499)
First time entry in checklist	0.523	0.582	0.483	0.514
(2011 and after)	(0.500)	(0.494)	(0.500)	(0.500)
Observations	3695	1102	1306	1287

Table 3.3. Descriptive Statistics for Dependent Variables

	(1)	(2)	(3)	(4)
	Full sample	**Inexperienced**	**1-3-year experience**	**More than 3-year experience**
WWPs				
Health risk appraisals	0.669	0.411	0.761	0.797
	(0.471)	(0.492)	(0.427)	(0.402)
Health screening	0.697	0.425	0.771	0.855
	(0.460)	(0.495)	(0.420)	(0.352)
Physical activity	0.835	0.575	0.938	0.954
	(0.371)	(0.495)	(0.241)	(0.209)
Smoking cessation	0.673	0.447	0.732	0.805
	(0.469)	(0.497)	(0.443)	(0.396)
Nutrition/weight Management	0.821	0.549	0.931	0.942
	(0.384)	(0.498)	(0.253)	(0.234)
Responsible alcohol use	0.355	0.206	0.380	0.457
	(0.479)	(0.405)	(0.486)	(0.498)
Stress management	0.701	0.422	0.786	0.855
	(0.458)	(0.494)	(0.411)	(0.353)
Medical self-care	0.490	0.265	0.523	0.650
	(0.500)	(0.442)	(0.500)	(0.477)
Work family	0.432	0.235	0.467	0.564
	(0.495)	(0.424)	(0.499)	(0.496)
Personal financial management	0.524	0.340	0.543	0.663
	(0.499)	(0.474)	(0.498)	(0.473)
Safety/health protection	0.705	0.535	0.733	0.822
	(0.456)	(0.499)	(0.443)	(0.383)
Ergonomics	0.481	0.319	0.505	0.598
	(0.500)	(0.466)	(0.500)	(0.491)
Mental health/depression	0.549	0.311	0.595	0.707
	(0.498)	(0.463)	(0.491)	(0.455)
Disease management	0.621	0.366	0.688	0.772
	(0.485)	(0.482)	(0.463)	(0.419)
Observations	3695	1102	1306	1287

Table 3.4. Multivariate Probit Results on Inexperienced Sample

Companies with no experience in wellness programs (N = 1102)	**Health risk appraisal**	**Health screening**	**Physical activity**	**Smoking cessation**	**Nutrition/ weight**	**Alcohol use**	**Stress Management**
Improve employee health	-0.252 (0.244)	0.095 (0.262)	0.271 (0.265)	-0.061 (0.269)	0.236 (0.254)	0.284 (0.291)	0.204 (0.272)
Improve the health of spouses and dependents	-0.164 (0.274)	0.111 (0.293)	0.207 (0.299)	0.010 (0.299)	0.331 (0.290)	0.455 (0.327)	0.264 (0.305)
Increase responsibility for managing personal health	0.057 (0.254)	**0.590**** **(0.271)**	0.264 (0.277)	-0.005 (0.280)	0.310 (0.266)	0.410 (0.303)	0.344 (0.283)
Contain health care costs	0.169 (0.243)	**0.478*** **(0.261)**	0.227 (0.268)	0.085 (0.272)	0.126 (0.257)	0.346 (0.293)	0.388 (0.275)
Produce ROI	-0.024 (0.289)	0.368 (0.305)	0.170 (0.303)	-0.336 (0.313)	0.032 (0.291)	0.102 (0.339)	0.186 (0.314)
Reduce unnecessary medical care utilization	-0.289 (0.276)	-0.013 (0.290)	-0.193 (0.296)	-0.254 (0.303)	-0.263 (0.282)	0.103 (0.327)	0.013 (0.300)
Improve performance or Productivity or reduce absenteeism	-0.433 (0.291)	0.002 (0.302)	0.044 (0.308)	-0.197 (0.315)	-0.085 (0.300)	0.340 (0.340)	0.224 (0.313)
Retention or employee request	-0.334 (0.272)	0.193 (0.286)	0.348 (0.296)	-0.123 (0.298)	0.083 (0.282)	0.156 (0.318)	0.485 (0.301)
Improve employee morale	-0.258 (0.255)	0.171 (0.273)	0.179 (0.277)	-0.049 (0.283)	0.161 (0.269)	0.292 (0.304)	0.341 (0.283)
Industry: Services	-0.051 (0.088)	-0.080 (0.089)	0.018 (0.087)	0.083 (0.088)	0.012 (0.086)	-0.123 (0.095)	0.047 (0.084)
Industry: Manufacturing	-0.032 (0.122)	0.092 (0.120)	-0.100 (0.123)	0.382*** (0.124)	0.105 (0.123)	0.142 (0.131)	-0.053 (0.118)
Small company	-0.217** (0.087)	-0.343*** (0.086)	-0.205** (0.084)	-0.541*** (0.086)	-0.337*** (0.084)	-0.164* (0.092)	-0.213*** (0.082)
The first year of entry to the Checklist: 2011 and after	0.018 (0.078)	-0.050 (0.079)	-0.130* (0.078)	-0.055 (0.078)	-0.062 (0.077)	-0.026 (0.086)	-0.045 (0.076)
Constant	0.185 (0.478)	-0.376 (0.511)	-0.025 (0.526)	0.260 (0.535)	0.056 (0.504)	-1.247** (0.578)	-0.613 (0.539)

Standard errors are in parenthesis. * p<0.10, ** p<0.05, *** p<0.01.

Table 3.4. Multivariate Probit Results on Inexperienced Sample (Continued)

Companies with no experience in wellness programs (N = 1102)	**Medical self-care**	**Work family**	**Personal financial**	**Safety/ health protection**	**Ergonomics**	**Mental health/ depression**	**Disease management**
Improve employee health	0.016 (0.280)	0.142 (0.298)	0.318 (0.275)	0.221 (0.261)	0.022 (0.271)	-0.140 (0.267)	-0.018 (0.277)
Improve the health of spouses and dependents	0.229 (0.313)	0.299 (0.332)	**0.500* (0.303)**	0.279 (0.292)	0.015 (0.303)	0.094 (0.298)	0.173 (0.307)
Increase responsibility for managing personal health	0.335 (0.293)	**0.535* (0.308)**	**0.644** (0.284)**	0.377 (0.272)	-0.021 (0.282)	0.171 (0.279)	0.267 (0.287)
Contain health care costs	0.211 (0.284)	0.355 (0.298)	0.423 (0.275)	0.342 (0.263)	-0.090 (0.274)	0.038 (0.269)	0.071 (0.278)
Produce ROI	0.030 (0.322)	0.217 (0.341)	0.433 (0.315)	0.200 (0.304)	-0.012 (0.317)	-0.141 (0.305)	0.094 (0.312)
Reduce unnecessary medical care utilization	-0.202 (0.318)	0.018 (0.330)	-0.075 (0.311)	0.132 (0.289)	-0.263 (0.302)	-0.206 (0.300)	-0.183 (0.306)
Improve performance or Productivity or reduce absenteeism	0.112 (0.329)	0.322 (0.333)	0.293 (0.317)	0.266 (0.306)	-0.140 (0.320)	-0.094 (0.313)	-0.049 (0.315)
Retention or employee request	0.246 (0.312)	0.422 (0.323)	0.386 (0.293)	0.144 (0.286)	-0.254 (0.304)	0.045 (0.294)	0.107 (0.304)
Improve employee morale	0.176 (0.296)	0.349 (0.309)	**0.485* (0.286)**	0.285 (0.274)	-0.062 (0.285)	-0.201 (0.284)	-0.002 (0.291)
Industry: Services	-0.024 (0.091)	-0.004 (0.092)	-0.006 (0.089)	-0.153* (0.085)	0.121 (0.092)	0.053 (0.090)	0.021 (0.087)
Industry: Manufacturing	0.124 (0.126)	0.013 (0.130)	-0.054 (0.124)	0.518*** (0.123)	0.680*** (0.123)	-0.008 (0.125)	0.057 (0.122)
Small company	-0.115 (0.087)	-0.208** (0.088)	-0.199** (0.085)	-0.069 (0.085)	-0.126 (0.089)	-0.301*** (0.085)	-0.323*** (0.084)
The first year of entry to the Checklist: 2011 and after	-0.116 (0.081)	-0.047 (0.082)	-0.079 (0.079)	-0.104 (0.078)	-0.188** (0.081)	-0.143* (0.081)	-0.070 (0.078)
Constant	-0.733 (0.558)	-1.102 (0.592)	-0.984* (0.541)	-0.328 (0.517)	-0.344 (0.538)	-0.133 (0.530)	-0.178 (0.549)

Standard errors are in parenthesis. * p<0.10, ** p<0.05, *** p<0.01

Table 3.5. Multivariate Probit Results on 1-3-Year Experienced Sample

Companies with 1-3-year experience in wellness programs (N = 1306)	**Health risk appraisal**	**Health screening**	**Physical activity**	**Smoking cessation**	**Nutrition/ weight**	**Alcohol use**	**Stress Management**
Improve employee health	-0.226 (0.416)	-0.047 (0.425)	-0.197 (0.699)	**0.642*** **(0.385)**	0.025 (0.523)	-0.005 (0.423)	-0.622 (0.498)
Improve the health of spouses and dependents	-0.032 (0.435)	-0.010 (0.441)	-0.132 (0.719)	**0.691*** **(0.403)**	-0.035 (0.545)	0.124 (0.437)	-0.438 (0.509)
Increase responsibility for managing personal health	-0.095 (0.421)	-0.103 (0.429)	-0.239 (0.700)	0.419 (0.390)	0.043 (0.530)	-0.009 (0.427)	-0.644 (0.500)
Contain health care costs	0.181 (0.421)	0.016 (0.428)	-0.291 (0.693)	**0.720*** **(0.389)**	0.117 (0.525)	0.081 (0.425)	-0.448 (0.497)
Produce ROI	0.138 (0.455)	0.189 (0.460)	-0.105 (0.738)	**0.996**** **(0.427)**	0.127 (0.569)	0.225 (0.450)	-0.393 (0.529)
Reduce unnecessary medical care utilization	0.023 (0.445)	-0.107 (0.450)	-0.575 (0.719)	0.595 (0.417)	-0.123 (0.564)	0.055 (0.447)	-0.689 (0.519)
Improve performance or Productivity or reduce absenteeism	-0.371 (0.436)	-0.316 (0.445)	-0.534 (0.731)	**0.750*** **(0.417)**	-0.364 (0.550)	0.163 (0.447)	-0.464 (0.513)
Retention or employee request	-0.115 (0.443)	-0.159 (0.448)	-0.144 (0.744)	**0.826**** **(0.418)**	0.308 (0.578)	0.066 (0.448)	-0.407 (0.527)
Improve employee morale	-0.371 (0.430)	-0.400 (0.438)	-0.038 (0.714)	0.511 (0.399)	0.117 (0.542)	0.122 (0.436)	-0.447 (0.505)
Industry: Services	-0.118 (0.086)	-0.057 (0.085)	-0.204* (0.123)	0.052 (0.083)	-0.283** (0.117)	-0.065 (0.079)	-0.106 (0.086)
Industry: Manufacturing	0.109 (0.120)	0.020 (0.120)	-0.418*** (0.153)	0.349*** (0.117)	-0.220 (0.158)	0.153 (0.105)	-0.320*** (0.114)
Small company	-0.260*** (0.079)	-0.149* (0.078)	-0.088 (0.104)	-0.391*** (0.078)	-0.244** (0.104)	-0.067 (0.072)	-0.234*** (0.079)
The first year of entry to the Checklist: 2011 and after	0.021 (0.077)	-0.141* (0.077)	0.067 (0.104)	-0.094 (0.076)	-0.101 (0.102)	-0.181** (0.071)	-0.103 (0.076)
Constant	1.049 (0.827)	1.057 (0.843)	2.244 (1.387)	-0.483 (0.766)	1.787 (1.038)	-0.273 (0.843)	2.146** (0.990)

Standard errors are in parenthesis. * p<0.10, ** p<0.05, *** p<0.01

Table 3.5. Multivariate Probit Results on 1-3-Year Experienced Sample (Continued)

Companies with 1-3-year experience in wellness programs (N = 1306)	**Medical self-care**	**Work family**	**Personal financial**	**Safety/ health protection**	**Ergonomics**	**Mental health/ depression**	**Disease management**
Improve employee health	-0.690 (0.483)	-0.043 (0.406)	-0.317 (0.411)	-0.464 (0.439)	**-1.011***** **(0.390)**	-0.273 (0.409)	-0.391 (0.429)
Improve the health of spouses and dependents	-0.299 (0.495)	0.373 (0.423)	-0.019 (0.427)	-0.354 (0.453)	**-0.919**** **(0.403)**	0.085 (0.424)	-0.041 (0.442)
Increase responsibility for managing personal health	-0.658 (0.487)	-0.182 (0.411)	-0.412 (0.415)	-0.449 (0.442)	**-0.986**** **(0.395)**	-0.291 (0.413)	-0.364 (0.434)
Contain health care costs	-0.576 (0.485)	-0.030 (0.409)	-0.293 (0.414)	-0.314 (0.437)	**-0.942**** **(0.397)**	-0.130 (0.411)	-0.250 (0.432)
Produce ROI	-0.461 (0.507)	0.119 (0.435)	-0.272 (0.438)	-0.263 (0.467)	**-0.791*** **(0.421)**	-0.094 (0.433)	-0.182 (0.462)
Reduce unnecessary medical care utilization	-0.723 (0.503)	-0.238 (0.434)	-0.407 (0.436)	-0.424 (0.465)	**-1.044**** **(0.417)**	-0.043 (0.437)	-0.251 (0.454)
Improve performance or Productivity or reduce absenteeism	-0.630 (0.504)	0.047 (0.431)	-0.265 (0.437)	-0.107 (0.472)	**-0.806*** **(0.417)**	-0.240 (0.434)	-0.145 (0.453)
Retention or employee request	-0.648 (0.507)	0.136 (0.431)	-0.190 (0.435)	-0.446 (0.460)	-0.468 (0.411)	0.113 (0.436)	-0.239 (0.453)
Improve employee morale	-0.561 (0.493)	0.104 (0.418)	-0.298 (0.423)	-0.346 (0.448)	**-0.920**** **(0.405)**	-0.027 (0.422)	-0.389 (0.440)
Industry: Services	-0.118 (0.078)	-0.092 (0.077)	-0.020 (0.078)	-0.177** (0.082)	0.149* (0.078)	-0.067 (0.079)	-0.073 (0.082)
Industry: Manufacturing	-0.068 (0.106)	-0.191* (0.109)	-0.170 (0.108)	0.307** (0.122)	0.439*** (0.107)	-0.109 (0.108)	0.005 (0.111)
Small company	0.008 (0.071)	-0.303*** (0.072)	-0.237*** (0.072)	-0.126* (0.076)	-0.135 (0.072)	-0.324*** (0.073)	-0.479*** (0.076)
The first year of entry to the Checklist: 2011 and after	-0.190*** (0.070)	-0.165** (0.071)	-0.066 (0.071)	-0.092 (0.074)	-0.096 (0.070)	-0.158** (0.071)	-0.096 (0.073)
Constant	1.444 (0.964)	0.280 (0.811)	0.898 (0.819)	1.557* (0.871)	1.869** (0.777)	0.901 (0.814)	1.461* (0.852)

Standard errors are in parenthesis. * p<0.10, ** p<0.05, *** p<0.01

Table 3.6. Multivariate Probit Results on More Than 3-Year Experienced Sample

Companies with more than 3-year experience in wellness programs (N = 1287)	**Health risk appraisal**	**Health screening**	**Physical activity**	**Smoking cessation**	**Nutrition/ weight**	**Alcohol use**	**Stress Management**
Improve employee health	0.347 (0.325)	0.370 (0.319)	**1.115*** (0.368)**	0.370 (0.321)	0.333 (0.392)	0.269 (0.285)	**-0.812** (0.414)**
Improve the health of spouses and dependents	**0.694** (0.347)**	**0.867** (0.359)**	**1.000** (0.393)**	**0.888** (0.351)**	0.202 (0.407)	**0.613** (0.301)**	-0.611 (0.427)
Increase responsibility for managing personal health	0.388 (0.332)	0.353 (0.331)	**0.887** (0.363)**	0.468 (0.329)	0.118 (0.397)	0.232 (0.292)	**-0.794* (0.421)**
Contain health care costs	0.598 (0.325)	0.341 (0.324)	**0.679* (0.354)**	0.341 (0.323)	0.253 (0.379)	0.074 (0.287)	**-0.867** (0.420)**
Produce ROI	**0.359* (0.388)**	0.170 (0.389)	**0.920* (0.482)**	0.367 (0.389)	-0.110 (0.464)	-0.142 (0.342)	**-1.082** (0.464)**
Reduce unnecessary medical care utilization	0.442 (0.357)	0.559 (0.361)	0.600 (0.393)	0.399 (0.355)	-0.129 (0.420)	-0.031 (0.315)	**-0.965** (0.443)**
Improve performance or Productivity or reduce absenteeism	0.185 (0.364)	0.216 (0.362)	0.610 (0.413)	0.565 (0.361)	0.178 (0.436)	0.473 (0.321)	-0.694 (0.436)
Retention or employee request	0.304 (0.352)	0.363 (0.352)	**1.164** (0.468)**	0.411 (0.360)	0.009 (0.447)	0.109 (0.322)	**-0.919** (0.452)**
Improve employee morale	-0.029 (0.343)	-0.046 (0.343)	**0.666* (0.374)**	0.187 (0.343)	0.300 (0.415)	0.002 (0.307)	-0.610 (0.441)
Industry: Services	-0.049 (0.090)	-0.015 (0.094)	0.083 (0.127)	0.104 (0.087)	0.099 (0.117)	-0.073 (0.078)	0.190** (0.093)
Industry: Manufacturing	0.101 (0.125)	0.140 (0.134)	-0.207 (0.152)	0.393*** (0.130)	0.275 (0.181)	0.050 (0.104)	0.032 (0.122)
Small company	-0.186** (0.081)	-0.252*** (0.089)	-0.291** (0.114)	-0.556*** (0.082)	-0.368*** (0.110)	-0.169** (0.072)	-0.326*** (0.087)
The first year of entry to the Checklist: 2011 and after	-0.004 (0.081)	-0.030 (0.087)	0.012 (0.117)	0.103 (0.082)	0.045 (0.103)	-0.123* (0.071)	-0.071 (0.083)
Constant	0.102 (0.639)	0.475 (0.630)	0.081 (0.700)	0.193 (0.633)	1.187 (0.756)	-0.355 (0.562)	2.773*** (0.821)

Standard errors are in parenthesis. * p<0.10, ** p<0.05, *** p<0.01

Table 3.6. Multivariate Probit Results on More Than 3-Year Experienced Sample (Continued)

Companies with more than 3-year experience in wellness programs (N = 1287)	**Medical self-care**	**Work family**	**Personal financial**	**Safety/ health protection**	**Ergonomics**	**Mental health/ depression**	**Disease management**
Improve employee health	0.165 (0.289)	-0.083 (0.277)	-0.187 (0.302)	0.200 (0.318)	0.323 (0.286)	-0.185 (0.290)	0.234 (0.313)
Improve the health of spouses and dependents	0.448 (0.308)	0.275 (0.294)	-0.086 (0.313)	0.540 (0.338)	**0.603** (0.304)**	0.157 (0.308)	**0.639** (0.321)**
Increase responsibility for managing personal health	0.275 (0.296)	0.052 (0.283)	-0.113 (0.311)	0.380 (0.324)	0.460 (0.293)	-0.059 (0.296)	0.190 (0.323)
Contain health care costs	0.037 (0.289)	-0.022 (0.275)	-0.170 (0.305)	0.124 (0.315)	0.226 (0.286)	-0.167 (0.291)	0.077 (0.318)
Produce ROI	-0.296 (0.340)	-0.088 (0.338)	-0.351 (0.357)	0.034 (0.373)	0.153 (0.343)	-0.332 (0.344)	0.107 (0.372)
Reduce unnecessary medical care utilization	0.065 (0.316)	-0.273 (0.304)	-0.430 (0.330)	0.169 (0.341)	0.312 (0.314)	0.018 (0.325)	0.243 (0.349)
Improve performance or Productivity or reduce absenteeism	0.312 (0.330)	0.146 (0.310)	-0.122 (0.339)	0.411 (0.358)	**0.675** (0.327)**	-0.010 (0.330)	0.088 (0.365)
Retention or employee request	0.146 (0.325)	0.057 (0.313)	-0.158 (0.337)	0.261 (0.352)	0.450 (0.319)	-0.043 (0.330)	0.045 (0.353)
Improve employee morale	-0.135 (0.308)	-0.111 (0.295)	-0.183 (0.324)	0.085 (0.336)	0.172 (0.308)	-0.465 (0.312)	-0.015 (0.337)
Industry: Services	-0.056 (0.081)	0.011 (0.079)	0.047 (0.081)	-0.069 (0.089)	0.011 (0.080)	0.217*** (0.082)	-0.178** (0.084)
Industry: Manufacturing	-0.011 (0.106)	-0.078 (0.105)	0.033 (0.108)	0.410*** (0.133)	0.060 (0.105)	0.033 (0.111)	0.044 (0.121)
Small company	-0.195** (0.074)	-0.413*** (0.073)	-0.319*** (0.074)	-0.205** (0.083)	-0.267*** (0.074)	-0.372*** (0.077)	-0.485*** (0.079)
The first year of entry to the Checklist: 2011 and after	-0.055 (0.073)	0.095 (0.071)	0.059 (0.073)	-0.317*** (0.083)	-0.041 (0.072)	0.067 (0.074)	-0.011 (0.078)
Constant	0.250 (0.566)	0.339 (0.540)	0.851 (0.597)	0.725 (0.623)	-0.298 (0.562)	0.855 (0.569)	0.700 (0.621)

Standard errors are in parenthesis. * p<0.10, ** p<0.05, *** p<0.01

References

1. Anderko L, Roffenbender JS, Goetzel RZ, Millard F, Wildenhaus K, et al. (2012). Promoting Prevention through the Affordable Care Act: Workplace Wellness. Prev Chronic Dis, 9.

2. Baicker, K., Cutler, D., & Song, Z. (2010). Workplace wellness programs can generate savings. *Health affairs*, 29(2), 304-311.

3. Baker, K. M., Goetzel, R. Z., Pei, X., Weiss, A. J., Bowen, J., Tabrizi, M. J. & Thompson, E. (2008). Using a return-on-investment estimation model to evaluate outcomes from an obesity management worksite health promotion program. *Journal of Occupational and Environmental Medicine*, 50(9), 981-990.

4. Barbosa, C., Bray, J.W., Dowd, W.N., Mills, M.J., Kelly, E.L., Moen, P., & Wipfli, B. (2015). Return on investment of a Work-Family Intervention: Evidence from the Work, Family, and Health Network. *Journal of Occupational and Environmental Medicine*, 57(9), 943–951.

5. Benjamin, G. C. (2006). Putting the public in public health: New approaches. *Health Affairs*, *25*(4), 1040-1043.

6. Berry, L., Mirabito, A., & Baun, W. (2010). What's the hard return on employee wellness programs?

7. Bertera, R. L. (1990). The effects of workplace health promotion on absenteeism and employment costs in a large industrial population. *American Journal of Public Health*, 80(9), 1101-1105.

8. Cappellari, L., & Jenkins, S. P. (2003). Multivariate probit regression using simulated maximum likelihood. *The Stata Journal*, *3*(3), 278-294.

9. Chapman, L. S. (2004). Expert opinions on" best practices" in worksite health promotion (WHP). *American journal of health promotion: AJHP*, *18*(6), 1-6.

10. Declaration, J. (1997). Jakarta Declaration on Leading Health Promotion into the 21st Century. *Fourth International Conference on Health Promotion, Jakarta*.

11. Edington, D. W., Pitts, J. S., & Schultz, A. B. (2014). Health as an economic strategy: special emphasis on the workplace. *American Journal of Lifestyle Medicine*, *8*(5), 330-335.

12. Edington, D. W., Schultz, A. B., Pitts, J. S., & Camilleri, A. (2016). The Future of Health Promotion in the 21st Century: A Focus on the Working Population. *American Journal of Lifestyle Medicine*, *10*(4), 242-252.

13. Fielding, J. E. (1984). Health promotion and disease prevention at the worksite. *Annual Review of Public Health*, *5*(1), 237-265.

14. Goetzel, R. Z., Guindon, A., Humphries, L., Newton, P., Turshen, J., & Webb, R. (1998). Health and productivity management: consortium benchmarking study best practice report. *Houston, TX: American Productivity and Quality Center International Benchmarking Clearinghouse*.

15. Guazzi, M., Faggiano, P., Mureddu, G. F., Faden, G., Niebauer, J., & Temporelli, P. L. (2014). Worksite health and wellness in the European Union. *Progress in cardiovascular diseases*, *56*(5), 508-514.

16. Hunnicutt, D; Leffelman, B. (2007). WELCOA's Well Workplace Initiative. *WELCOA's Absolute Advantage Magazine,* 6(10), 2-3.

17. Kaiser Family Foundation, Health Research and Educational Trust. (2017). Employer health benefits: 2017 annual survey.

18. Kent, K., Goetzel, R. Z., Roemer, E. C., Prasad, A., & Freundlich, N. (2016). Promoting healthy workplaces by building cultures of health and applying strategic communications. *Journal of occupational and environmental medicine*, *58*(2), 114-122.

19. Khoury, A. (2014). The Evolution of Worksite Wellness. Corporate Wellness Magazine. Accessed May 23, 2017 from http://www.corporatewellnessmagazine.com/worksite-wellness/the-evolution-of/.

20. Kickbusch, I., & Payne, L. (2003). Twenty-first century health promotion: the public health revolution meets the wellness revolution. *Health Promotion International*, 18 (4): 275-278.

21. Kirkland, A. (2014). What Is Wellness Now?. *Journal of Health Politics*, Policy and Law, 39(5).

22. Lewis, A. and Khanna, V. S. (2013, January 16) Is It Time to Re-Examine Workplace Wellness 'Get Well Quick' Schemes? Accessed January 28, 2017, from http://thehealthcareblog.com/blog/2014/12/12/the-wellness-industrys-terrible-horrible-no-good-very-bad-we/

23. Lewis, A. and Khanna, V. S. (2014, December 12). The Wellness Industries Terrible, Horrible No Good Very Bad Week. Accessed May 28, 2017, from

http://healthaffairs.org/blog/2013/01/16/is-it-time-to-re-examine-workplace-wellness-get-well-quick-schemes/

24. Lewis, A., Khanna, V., and Montrose, S. (2014, November 25). Workplace Wellness Produces No Savings. Accessed January 28, 2017, from http://healthaffairs.org/blog/2014/11/25/workplace-wellness-produces-no-savings

25. Mattke, S., Liu, H., Caloyeras, J., Huang, C. Y., Van Busum, K. R., Khodyakov, D., & Shier, V. (2013). Workplace wellness programs study: Final report. *Rand Health Quarterly*, 3(2)

26. Martínez-Lemos, R. I. (2015). Economic impact of corporate wellness programs in Europe: A literature review. *Journal of occupational health*, *57*(3), 201-211.

27. Mas-Collel, A., Whinston, M.D., Green, J.R., 1995. Microeconomic Theory. Oxford University Press, New York.

28. National Institute for Occupational Health and Safety (NIOSH) (2012). *Research Compendium: The NIOSH Total Worker Health™ Program: Seminal Research Papers* 2012. Washington, DC: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, DHHS (NIOSH) Publication No. 2012-146:1-214. Accessed January 28, 2017, from https://www.cdc.gov/niosh/docs/2012-146/pdfs/2012-146.pdf

29. O'Donnell, M. P. (2016). Is It Time to Separate the Financial and Health Goals of Workplace Health Promotion Programs?. *American journal of health promotion: AJHP*, *30*(7), 496. Patient Protection and Affordable Care Act, 42 U.S.C. § 10408, (2010).

30. Patterson, M., Warr, P., & West, M. (2004). Organizational climate and company productivity: The role of employee affect and employee level. *Journal of Occupational and Organizational Psychology*, *77*(2), 193-216.

31. Schill, A. L., & Chosewood, L. C. (2013). The NIOSH Total Worker Health™ program: an overview.

32. Terry, P. E., Seaverson, E. L., Grossmeier, J., & Anderson, D. R. (2008). Association between nine quality components and superior worksite health management program results. *Journal of occupational and environmental medicine*, *50*(6), 633-641.

33. World Health Organization (WHO) (1997). WHO's Global Healthy Work Approach. Division of Health Promotion, Education & Communication and Office of Occupational Health, Geneva.

34. World Health Organization. WHO (2002) Reducing Risks, Promoting Healthy Life.

35. World Health Organization, Public Health Agency of Canada, & Canada. Public Health Agency of Canada. (2005). Preventing chronic diseases: a vital investment. World Health Organization.

36. World Health Organization. (2005). Participants at the 6th Global Conference on Health Promotion. The Bangkok Charter for health promotion in a globalized world. Geneva, Switzerland: World Health Organization.

Appendix 3.A. The First Order Conditions

This section sets up the optimization problem of a representative firm that chooses its labor supply and WWP to maximize its objective function. We use a random utility model to establish a functional relationship between program options, implementation reasons, and organizations' characteristics. The objective function has a systematic utility (V), and a random residual (e). The systematic utility is a function of profit (π) and other utility components (ϕ) that are not profit related.

$$u = V(\pi, \phi) + e \tag{3.A.1}$$

We assumed that there are three reasons for WWP implementation (R) and two WWPs ($x \epsilon \{ x_1, x_2 \}$) that the firm can choose to implement to achieve its objectives. Reasons for WWP implementation are savings from health care and productivity- related costs (S), improving employee health (H), and improving employee morale (M). x_1 and x_2 represent the two unspecified WWPs that a firm can chose between.

$$R \in \{S, H, M\}$$
$$x \in \{x_1, x_2\}$$

The program choice is assumed to attain the reason for program implementation. Thus, each reason for WWP implementation is a function of WWP choices:

$$S = s(x) \quad (3.A.2)$$
$$H = h(x)$$
$$M = m(x)$$

where *S, H, M* can be improved with the right program input *x*.

The profit component is the difference between total revenue and total cost. Total revenue is calculated as price *(p)* times production *(Y)*. Total cost is determined with cost function *(C)*.

$$\pi = pY - C \quad (3.A.3)$$

Production is the function of labor input (*L*). Labor is a function of employee health where better health increases labor productivity.

$$Y = f(L) \quad (3.A.4)$$

$$L = L(H) \quad (3.A.5)$$

where $f(.)$ is increasing in L at a decreasing rate. In addition, labor is a function of employee health where the productivity increases with better health inputs.

Cost function includes labor cost, workers' compensation claims, and program cost. Wages (*w*), health care benefits (*b*), and sick days (*d*) determine the labor cost. The labor cost is the sum of health care benefits (*bwL*), cost of lost days (absenteeism) (*dwL*), and cost of presenteeism (τwL). If there is no productivity loss due to presenteeism, then τ =1 and the cost is *wL*. If there is loss due to presenteeism, then τ >1 and the cost is

τwL. Worker's compensation claims (*comp*) are the cost of work-related injuries. If any program is chosen given the reason for program implementation, represented by the indicator function (*I(x|R)*), then the cost function includes program costs (pc_x).

$$C = (\tau(x) + b(x) + d(x))wL(h(x)) + \theta_1 comp(x) + \theta_2 pc_x\{I(x \mid R)\} \tag{3.A.6}$$

The utility components that are not profit related (ϕ) create a corporate citizenship value for organizations to achieve firm's goals (Peredo and McLean, 2006). Employee health and employee morale both affect this component, which is the additional piece to the classical firm theory of profit maximizing objectives. Including this component helps us to identify different motivations of for-profit, not-for profit, and nonprofit companies.

$$\phi = \phi(H, M) \tag{3.A.7}$$

Suppose that the objective function has the linear form with parameters α and β that are organization's weight on profit and other utility components. An organization maximizes its utility with respect to labor input and WWP choices:

$$\max_{x\in\{x_1,x_2\},L>0}\{\alpha\pi + \beta\phi + e\} \tag{3.A.8}$$

Plug equations (3.A.2) to (3.A.7) in to (3.A.8). Let $\tau(x) + b(x) + d(x) = s(x)$. An organization maximizes its utility with respect to labor input and WWP choices. Note that

choice of wellness programs is discrete. Normalize market price of production to 1 (p=1). The first order conditions (FOC) with respect to the choice variables:

$$\frac{\partial u}{\partial L} = \alpha\left(\frac{df(L)}{d(L)} - \frac{\partial C}{\partial L}\right) = 0 \tag{3.A.9}$$

$$\frac{df(L)}{d(L)} - \frac{\partial C}{\partial L} = f_L - s(x)w = 0$$

$$f_L = s(x)w$$

$$\frac{\Delta u}{\Delta x_n} = \alpha\left(\frac{\Delta f(L)}{\Delta x_n} - \frac{\Delta C}{\Delta x_n}\right) + \beta\left(\frac{\Delta\phi(H,M)}{\Delta x_n}\right) \tag{3.A.10}$$

$$= \alpha\left(\frac{\Delta f(L)}{\Delta L}\frac{\Delta L}{\Delta H}\frac{\Delta h(x)}{\Delta x_n} - \frac{\Delta s(x)}{\Delta x_n}wL(h(x)) - s(x)w\frac{\Delta L}{\Delta H}\frac{\Delta h(x)}{\Delta x_n} - \theta_1\frac{\Delta comp(x)}{\Delta x_n} - \theta_2 pc_n\right) + \beta\left(\frac{\Delta\phi}{\Delta H}\frac{\Delta h(x)}{\Delta x_n} - \frac{\Delta\phi}{\Delta M}\frac{\Delta m(x)}{\Delta x_n}\right) \quad \forall n = 1,2$$

The demand for wellness programs is derived using these first order conditions. Using implicit function theorem, we can analyze how the objectives for implementing WWPs affect the WWP demand (Mas-Collel et al., 1995):

$$\frac{\Delta x_n}{\Delta R} = -\frac{\Delta u/\Delta R}{\Delta u/\Delta x_n} \tag{3.A.11}$$

Appendix 3.B. Additional Tables

Table 3.B1. Univariate Probit Results on Inexperienced Sample

Companies with no experience in wellness programs (N = 1102)	**Health risk appraisal**	**Health screening**	**Physical activity**	**Smoking cessation**	**Nutrition/ weight**	**Alcohol use**	**Stress Management**
Improve employee health	-0.273 (0.262)	0.099 (0.274)	0.236 (0.254)	-0.087 (0.257)	0.210 (0.256)	0.160 (0.310)	0.143 (0.262)
Improve the health of spouses and dependents	-0.159 (0.293)	0.169 (0.304)	0.175 (0.286)	0.007 (0.289)	0.284 (0.288)	0.402 (0.340)	0.227 (0.293)
Increase responsibility for managing personal health	0.042 (0.271)	**0.607**** **(0.283)**	0.233 (0.265)	-0.032 (0.268)	0.248 (0.267)	0.325 (0.319)	0.253 (0.272)
Contain health care costs	0.171 (0.262)	**0.512*** **(0.274)**	0.204 (0.255)	0.083 (0.258)	0.113 (0.257)	0.278 (0.310)	0.377 (0.263)
Produce ROI	0.005 (0.303)	0.413 (0.314)	0.199 (0.298)	-0.325 (0.304)	0.087 (0.300)	0.059 (0.361)	0.207 (0.304)
Reduce unnecessary medical care utilization	-0.299 (0.293)	-0.007 (0.305)	-0.237 (0.286)	-0.271 (0.290)	-0.266 (0.288)	-0.014 (0.349)	0.005 (0.295)
Improve performance or Productivity or reduce absenteeism	-0.421 (0.302)	0.054 (0.311)	-0.015 (0.292)	-0.219 (0.297)	-0.105 (0.294)	0.252 (0.349)	0.209 (0.299)
Retention or employee request	-0.321 (0.291)	0.236 (0.300)	0.273 (0.283)	-0.156 (0.285)	0.022 (0.284)	0.057 (0.343)	0.436 (0.289)
Improve employee morale	-0.281 (0.274)	0.159 (0.285)	0.133 (0.266)	-0.077 (0.269)	0.116 (0.267)	0.210 (0.322)	0.249 (0.273)
Industry: Services	-0.052 (0.088)	-0.076 (0.088)	0.061 (0.087)	0.121 (0.088)	0.055 (0.087)	-0.105 (0.098)	0.079 (0.087)
Industry: Manufacturing	-0.023 (0.124)	0.100 (0.124)	-0.103 (0.122)	0.394*** (0.124)	0.113 (0.123)	0.150 (0.132)	-0.048 (0.123)
Small company	-0.220** (0.087)	-0.342*** (0.087)	-0.226*** (0.088)	-0.562*** (0.087)	-0.372*** (0.088)	-0.130 (0.096)	-0.238*** (0.086)
The first year of entry to the Checklist: 2011 and after	0.027 (0.080)	-0.045 (0.080)	-0.163** (0.079)	-0.074 (0.079)	-0.101 (0.079)	0.001 (0.088)	-0.052 (0.079)
Constant	0.196 (0.517)	-0.433 (0.542)	0.071 (0.503)	0.299 (0.509)	0.149 (0.507)	-1.120* (0.615)	-0.519 (0.519)

Table 3.B1. Univariate Probit Results on Inexperienced Sample (Continued)

Companies with no experience in wellness programs (N = 1102)	**Medical self-care**	**Work family**	**Personal financial**	**Safety/ health protection**	**Ergonomics**	**Mental health/ depression**	**Disease management**
Improve employee health	-0.114 (0.274)	0.063 (0.294)	0.242 (0.280)	0.174 (0.255)	-0.017 (0.267)	-0.189 (0.264)	-0.037 (0.263)
Improve the health of spouses and dependents	0.205 (0.305)	0.318 (0.325)	0.490 (0.310)	0.275 (0.287)	0.035 (0.299)	0.113 (0.295)	0.218 (0.293)
Increase responsibility for managing personal health	0.230 (0.283)	0.467 (0.303)	**0.564*** **(0.289)**	0.305 (0.266)	-0.067 (0.278)	0.105 (0.273)	0.227 (0.272)
Contain health care costs	0.150 (0.274)	0.305 (0.295)	0.373 (0.280)	0.304 (0.257)	-0.118 (0.268)	0.022 (0.264)	0.094 (0.263)
Produce ROI	-0.006 (0.321)	0.280 (0.338)	0.438 (0.320)	0.193 (0.298)	0.016 (0.310)	-0.090 (0.309)	0.164 (0.305)
Reduce unnecessary medical care utilization	-0.288 (0.314)	-0.002 (0.332)	-0.077 (0.316)	0.139 (0.288)	-0.287 (0.304)	-0.198 (0.298)	-0.140 (0.296)
Improve performance or Productivity or reduce absenteeism	0.051 (0.314)	0.318 (0.334)	0.299 (0.316)	0.252 (0.295)	-0.126 (0.307)	-0.096 (0.304)	-0.009 (0.303)
Retention or employee request	0.157 (0.302)	0.356 (0.323)	0.320 (0.307)	0.083 (0.283)	-0.285 (0.298)	-0.001 (0.293)	0.097 (0.290)
Improve employee morale	0.028 (0.286)	0.236 (0.306)	0.400 (0.290)	0.229 (0.267)	-0.089 (0.279)	-0.282 (0.277)	-0.044 (0.274)
Industry: Services	-0.021 (0.093)	-0.002 (0.095)	0.005 (0.089)	-0.127 (0.086)	0.120 (0.092)	0.083 (0.091)	0.056 (0.088)
Industry: Manufacturing	0.101 (0.129)	-0.012 (0.135)	-0.072 (0.126)	0.548*** (0.128)	0.684*** (0.124)	0.006 (0.129)	0.066 (0.125)
Small company	-0.110 (0.092)	-0.195** (0.093)	-0.203** (0.088)	-0.074 (0.087)	-0.139 (0.089)	-0.310*** (0.088)	-0.344*** (0.087)
The first year of entry to the Checklist: 2011 and after	-0.090 (0.084)	-0.016 (0.086)	-0.064 (0.081)	-0.135* (0.079)	-0.183** (0.082)	-0.135* (0.082)	-0.070 (0.080)
Constant	-0.576 (0.542)	-1.024* (0.583)	-0.879 (0.554)	-0.233 (0.506)	-0.280 (0.529)	-0.087 (0.520)	-0.186 (0.519)

Standard errors are in parenthesis. * p<0.10, ** p<0.05, *** p<0.01

Table 3.B2 Univariate Probit Results on 1-3-Year Experienced Sample

Companies with 1-3-year experience in wellness programs (N = 1306)	**Health risk appraisal**	**Health screening**	**Physical activity**	**Smoking cessation**	**Nutrition/ weight**	**Alcohol use**	**Stress Management**
Improve employee health	-0.208 (0.414)	0.095 (0.262)	0.271 (0.265)	-0.061 (0.269)	0.236 (0.254)	0.284 (0.291)	0.204 (0.272)
Improve the health of spouses and dependents	0.005 (0.434)	0.111 (0.293)	0.207 (0.299)	0.010 (0.299)	0.331 (0.290)	0.455 (0.327)	0.264 (0.305)
Increase responsibility for managing personal health	-0.059 (0.418)	**0.590**** **(0.271)**	0.264 (0.277)	-0.005 (0.280)	0.310 (0.266)	0.410 (0.303)	0.344 (0.283)
Contain health care costs	0.218 (0.416)	**0.478*** **(0.261)**	0.227 (0.268)	0.085 (0.272)	0.126 (0.257)	0.346 (0.293)	0.388 (0.275)
Produce ROI	0.167 (0.454)	0.368 (0.305)	0.170 (0.303)	-0.336 (0.313)	0.032 (0.291)	0.102 (0.339)	0.186 (0.314)
Reduce unnecessary medical care utilization	0.051 (0.444)	-0.013 (0.290)	-0.193 (0.296)	-0.254 (0.303)	-0.263 (0.282)	0.103 (0.327)	0.013 (0.300)
Improve performance or Productivity or reduce absenteeism	-0.324 (0.438)	0.002 (0.302)	0.044 (0.308)	-0.197 (0.315)	-0.085 (0.300)	0.340 (0.340)	0.224 (0.313)
Retention or employee request	-0.094 (0.439)	0.193 (0.286)	0.348 (0.296)	-0.123 (0.298)	0.083 (0.282)	0.156 (0.318)	0.485 (0.301)
Improve employee morale	-0.342 (0.425)	0.171 (0.273)	0.179 (0.277)	-0.049 (0.283)	0.161 (0.269)	0.292 (0.304)	0.341 (0.283)
Industry: Services	-0.115 (0.086)	-0.080 (0.089)	0.018 (0.087)	0.083 (0.088)	0.012 (0.086)	-0.123 (0.095)	0.047 (0.084)
Industry: Manufacturing	0.111 (0.124)	0.092 (0.120)	-0.100 (0.123)	0.382*** (0.124)	0.105 (0.123)	0.142 (0.131)	-0.053 (0.118)
Small company	-0.274*** (0.080)	-0.343*** (0.086)	-0.205** (0.084)	-0.541*** (0.086)	-0.337*** (0.084)	-0.164* (0.092)	-0.213*** (0.082)
The first year of entry to the Checklist: 2011 and after	0.017 (0.078)	-0.050 (0.079)	-0.130* (0.078)	-0.055 (0.078)	-0.062 (0.077)	-0.026 (0.086)	-0.045 (0.076)
Constant	1.001 (0.823)	-0.376 (0.511)	-0.025 (0.526)	0.260 (0.535)	0.056 (0.504)	-1.247** (0.578)	-0.613 (0.539)

Standard errors are in parenthesis. * p<0.10, ** p<0.05, *** p<0.01

Table 3.B2 Univariate Probit Results on 1-3-Year Experienced Sample (Continued)

Companies with 1-3-year experience in wellness programs (N = 1306)	**Medical self-care**	**Work family**	**Personal financial**	**Safety/ health protection**	**Ergonomics**	**Mental health/ depression**	**Disease management**
Improve employee health	0.016 (0.280)	0.142 (0.298)	0.318 (0.275)	0.221 (0.261)	0.022 (0.271)	-0.140 (0.267)	-0.018 (0.277)
Improve the health of spouses and dependents	0.229 (0.313)	0.299 (0.332)	**0.500*** **(0.303)**	0.279 (0.292)	0.015 (0.303)	0.094 (0.298)	0.173 (0.307)
Increase responsibility for managing personal health	0.335 (0.293)	**0.535*** **(0.308)**	**0.644**** **(0.284)**	0.377 (0.272)	-0.021 (0.282)	0.171 (0.279)	0.267 (0.287)
Contain health care costs	0.211 (0.284)	0.355 (0.298)	0.423 (0.275)	0.342 (0.263)	-0.090 (0.274)	0.038 (0.269)	0.071 (0.278)
Produce ROI	0.030 (0.322)	0.217 (0.341)	0.433 (0.315)	0.200 (0.304)	-0.012 (0.317)	-0.141 (0.305)	0.094 (0.312)
Reduce unnecessary medical care utilization	-0.202 (0.318)	0.018 (0.330)	-0.075 (0.311)	0.132 (0.289)	-0.263 (0.302)	-0.206 (0.300)	-0.183 (0.306)
Improve performance or Productivity or reduce absenteeism	0.112 (0.329)	0.322 (0.333)	0.293 (0.317)	0.266 (0.306)	-0.140 (0.320)	-0.094 (0.313)	-0.049 (0.315)
Retention or employee request	0.246 (0.312)	0.422 (0.323)	0.386 (0.293)	0.144 (0.286)	-0.254 (0.304)	0.045 (0.294)	0.107 (0.304)
Improve employee morale	0.176 (0.296)	0.349 (0.309)	**0.485*** **(0.286)**	0.285 (0.274)	-0.062 (0.285)	-0.201 (0.284)	-0.002 (0.291)
Industry: Services	-0.024 (0.091)	-0.004 (0.092)	-0.006 (0.089)	-0.153* (0.085)	0.121 (0.092)	0.053 (0.090)	0.021 (0.087)
Industry: Manufacturing	0.124 (0.126)	0.013 (0.130)	-0.054 (0.124)	0.518*** (0.123)	0.680*** (0.123)	-0.008 (0.125)	0.057 (0.122)
Small company	-0.115 (0.087)	-0.208** (0.088)	-0.199** (0.085)	-0.069 (0.085)	-0.126 (0.089)	-0.301*** (0.085)	-0.323*** (0.084)
The first year of entry to the Checklist: 2011 and after	-0.116 (0.081)	-0.047 (0.082)	-0.079 (0.079)	-0.104 (0.078)	-0.188** (0.081)	-0.143* (0.081)	-0.070 (0.078)
Constant	-0.733 (0.558)	-1.102 (0.592)	-0.984* (0.541)	-0.328 (0.517)	-0.344 (0.538)	-0.133 (0.530)	-0.178 (0.549)

Standard errors are in parenthesis. * p<0.10, ** p<0.05, *** p<0.01

Table 3.B3. Univariate Probit Results on More Than 3-Year Experienced Sample

Companies with more than 3-year experience in wellness programs (N = 1287)	**Health risk appraisal**	**Health screening**	**Physical activity**	**Smoking cessation**	**Nutrition/ weight**	**Alcohol use**	**Stress Management**
Improve employee health	0.356 (0.302)	0.446 (0.332)	**1.057*** (0.368)**	0.360 (0.332)	0.385 (0.400)	0.241 (0.300)	-0.603 (0.499)
Improve the health of spouses and dependents	**0.721** (0.328)**	**0.954*** (0.365)**	**0.883** (0.400)**	**0.857** (0.358)**	0.211 (0.426)	**0.570* (0.316)**	-0.424 (0.514)
Increase responsibility for managing personal health	0.390 (0.312)	0.442 (0.341)	**0.890** (0.388)**	0.467 (0.341)	0.130 (0.411)	0.255 (0.306)	-0.599 (0.506)
Contain health care costs	**0.609** (0.306)**	0.414 (0.334)	**0.652* (0.368)**	0.323 (0.334)	0.303 (0.406)	0.081 (0.300)	-0.644 (0.501)
Produce ROI	0.387 (0.371)	0.245 (0.399)	**1.037* (0.554)**	0.388 (0.401)	0.056 (0.492)	-0.088 (0.354)	-0.720 (0.548)
Reduce unnecessary medical care utilization	0.440 (0.339)	0.573 (0.376)	0.488 (0.407)	0.357 (0.364)	-0.145 (0.434)	-0.040 (0.327)	-0.756 (0.521)
Improve performance or Productivity or reduce absenteeism	0.156 (0.345)	0.215 (0.381)	0.508 (0.441)	0.488 (0.383)	0.102 (0.464)	0.504 (0.339)	-0.542 (0.543)
Retention or employee request	0.299 (0.332)	0.430 (0.365)	**1.079** (0.446)**	0.407 (0.363)	0.051 (0.439)	0.097 (0.324)	-0.681 (0.523)
Improve employee morale	-0.017 (0.323)	-0.008 (0.351)	0.491 (0.406)	0.166 (0.352)	0.294 (0.442)	0.047 (0.320)	-0.378 (0.522)
Industry: Services	-0.046 (0.090)	0.007 (0.097)	-0.017 (0.142)	0.074 (0.091)	0.059 (0.125)	-0.059 (0.079)	0.137 (0.098)
Industry: Manufacturing	0.123 (0.126)	0.147 (0.137)	-0.198 (0.174)	0.361*** (0.133)	0.271 (0.187)	0.066 (0.106)	0.003 (0.128)
Small company	-0.198** (0.082)	-0.265*** (0.090)	-0.277** (0.129)	-0.534*** (0.085)	-0.339*** (0.118)	-0.156** (0.072)	-0.309*** (0.089)
The first year of entry to the Checklist: 2011 and after	-0.005 (0.082)	-0.004 (0.089)	-0.013 (0.127)	0.074 (0.084)	0.016 (0.116)	-0.122* (0.071)	-0.056 (0.088)
Constant	0.090 (0.597)	0.320 (0.654)	0.258 (0.715)	0.249 (0.655)	1.168 (0.791)	-0.350 (0.592)	2.382** (0.995)

Standard errors are in parenthesis. * p<0.10, ** p<0.05, *** p<0.01

Table 3.B3. Multivariate Probit Results on More Than 3-Year Experienced Sample (Continued)

Companies with more than 3-year experience in wellness programs (N = 1287)	**Medical self-care**	**Work family**	**Personal financial**	**Safety/ health protection**	**Ergonomics**	**Mental health/ depression**	**Disease management**
Improve employee health	0.149 (0.303)	-0.141 (0.302)	-0.204 (0.317)	0.199 (0.342)	0.322 (0.298)	-0.219 (0.336)	0.234 (0.324)
Improve the health of spouses and dependents	0.421 (0.319)	0.187 (0.318)	-0.126 (0.334)	0.551 (0.364)	**0.557*** **(0.314)**	0.085 (0.353)	**0.639*** **(0.348)**
Increase responsibility for managing personal health	0.292 (0.310)	0.036 (0.309)	-0.117 (0.324)	0.372 (0.352)	0.475 (0.305)	-0.079 (0.344)	0.197 (0.331)
Contain health care costs	0.048 (0.304)	-0.052 (0.304)	-0.180 (0.318)	0.138 (0.345)	0.224 (0.298)	-0.216 (0.338)	0.074 (0.325)
Produce ROI	-0.282 (0.354)	-0.068 (0.356)	-0.333 (0.369)	0.103 (0.404)	0.179 (0.351)	-0.331 (0.389)	0.129 (0.388)
Reduce unnecessary medical care utilization	0.079 (0.330)	-0.324 (0.329)	-0.439 (0.342)	0.191 (0.375)	0.339 (0.325)	0.027 (0.366)	0.275 (0.357)
Improve performance or Productivity or reduce absenteeism	0.280 (0.343)	0.100 (0.342)	-0.148 (0.356)	0.376 (0.393)	0.650* (0.339)	-0.059 (0.376)	0.084 (0.366)
Retention or employee request	0.141 (0.329)	0.010 (0.328)	-0.182 (0.342)	0.244 (0.376)	0.442 (0.324)	-0.066 (0.362)	0.052 (0.351)
Improve employee morale	-0.110 (0.322)	-0.099 (0.323)	-0.155 (0.337)	0.146 (0.365)	0.214 (0.317)	-0.487 (0.355)	0.017 (0.345)
Industry: Services	-0.078 (0.081)	0.015 (0.080)	0.026 (0.081)	-0.145 (0.092)	0.002 (0.080)	0.177** (0.084)	-0.222** (0.089)
Industry: Manufacturing	-0.001 (0.109)	-0.075 (0.107)	0.029 (0.109)	0.388*** (0.142)	0.058 (0.108)	0.008 (0.111)	0.009 (0.123)
Small company	-0.185** (0.074)	-0.421*** (0.072)	-0.318*** (0.074)	-0.153* (0.085)	-0.274*** (0.073)	-0.377*** (0.076)	-0.493*** (0.081)
The first year of entry to the Checklist: 2011 and after	-0.055 (0.073)	0.106 (0.072)	0.065 (0.073)	-0.323*** (0.085)	-0.030 (0.072)	0.083 (0.076)	-0.040 (0.080)
Constant	0.273 (0.598)	0.428 (0.597)	0.896 (0.627)	0.755 (0.676)	-0.283 (0.587)	0.958 (0.666)	0.765 (0.639)

Standard errors are in parenthesis. * p<0.10, ** p<0.05, *** p<0.01

CHAPTER IV

RETURN ON INVESTMENT OF WORKPLACE WELLNESS: EVIDENCE FROM A LONG-TERM CARE COMPANY

Introduction

For nearly half a century, businesses have relied on workplace wellness programs (WWPs) to improve employees' overall health and wellbeing, thereby reducing organizational costs such as health care costs, turnover cost, and productivity losses due to poor employee health (Cohen, 1985; Baicker et al., 2010; Neira, 2013). Business's expectations have been supported by an extensive body of empirical research, suggesting that WWPs improve employee health by reducing modifiable risks, such as obesity, tobacco use, physical inactivity, high stress, high blood pressure, and high blood glucose (Baker et al., 2008; Goetzel and Ozminowski, 2008; Meenan et al., 2010; Henke et al., 2011; Goetzel et al., 2012; Nyman et al., 2012; Goetzel et al., 2014; Musich et al., 2014; Widmer et al., 2014; Newman et al., 2015). Research has further linked these health improvements to decreased organizational costs associated with health care utilization, employee performance, and employee turnover (Baker et al., 2008; Goetzel and Ozminowski, 2008; Trogdon et al., 2009; Meenan et al., 2010; Yen et al., 2010; Henke et al., 2011; Nyman et al., 2012; Merrill et al., 2013; Musich et al., 2014; Barbosa et al., 2015; Dement et al., 2015). As a result, WWPs are commonplace in large companies in the US and worldwide.

In addition to the business community, WWPs are popular among governments and international health policy organizations as a public health tool. For example, the European Network for Workplace Health Promotion was launched in 1996 to increase the awareness of valuing and nurturing employees as well as to link WWPs with organizations' corporate strategies (ENWHP, 1997; De Greef, M., & Van den Broek, K., 2004). The World Health Organization engages governments, nongovernmental organizations, and the private sector to promote health using WWPs (Declaration, 1997). In the US, the federal government initiated the Prevention and Public Health Fund to strengthen public health via workplace wellness initiatives as part of the Affordable Care Act (ACA) of 2010 (Anderko et al., 2012). The ACA uses WWPs as a part of national public health policy to deal with chronic illnesses by giving grants to small businesses to provide comprehensive WWPs (Patient Protection and Affordable Care Act of 2010). Given public policy efforts worldwide that encourage private investment in WWPs, it is critical that we have a robust base of economic evaluation evidence supporting investment in WWPs.

The WWP economic evaluation literature has primarily focused on providing estimates of the return on investment (ROI) of WWPs because financial savings are considered to be the primary determinant WWP adoption by businesses (Hunnicutt and Leffelman, 2007; Meenan et al., 2010; Horwitz, Kelly, and Dinardo, 2013, Spence, 2015). Many studies find that WWPs are associated with positive ROI through reduced organizational costs associated with health care utilization, employee turnover, and productivity (Baker et al., 2008; Trogdon et al., 2009; Baicker et al., 2010; Yen et al.,

2010; Henke et al., 2011; Merrill et al., 2013; Goetzel et al., 2014; Dement et al., 2015). However, these findings on the ROI of WWPs are controversial. A growing number of studies show no indication that WWPs save money (Meenan et al., 2010; Horwitz et al., 2013; Lewis and Khanna, 2013; Lewis et al., 2014). Furthermore, WWP ROI research is often criticized for methodological flaws in the study design, such as randomization problems resulting from legal and practical issues, measurement inconsistencies such as different formulations of ROI, and improper statistical analysis such as a lack of statistical inference information on the ROI estimate (Lewis and Khanna, 2013, 2014; Lewis, Khanna, and Montrose, 2014). Emphasizing the importance of these criticisms, a systematic review on WWP ROI found that studies with higher methodological rigor had lower ROI estimates (Baxter, 2014).

In this study, we evaluate a WWP that was implemented by the Well-Spring Retirement Community in Greensboro, North Carolina. With approximately 300 employees, Well-Spring is a nonprofit organization meeting the US Small Business Administration's definition of a small business in Title 13, Code of Federal Regulation, part 121. Thus, our evaluation of the Well-Spring WWP is one of the first to be done of a small, nonprofit company. Beyond estimating the ROI of WWP in a small business, this paper contributes to the WWP ROI literature by presenting confidence intervals for our ROI estimates to allow for statistical inference. In addition to presenting an overall estimate of ROI, we estimate the effect of the WWP on turnover, employee injuries, lost employee labor days due to injury, and total organizational costs. We used the difference-in-differences method to address biases resulting from self-selection into program

participation. Findings show that WWP participants had significantly fewer lost labor days and lower organizational costs outcomes compared to non-participants. Overall, the financial savings from the WWP were higher than the program implementation costs, resulting in a statistically insignificant ROI of 0.59. However, the point estimate was not significant because of large confidence intervals.

Well-Spring's Wellness Program

Well-Spring, a nonprofit organization started in 1993, provides services addressing the physical, emotional, and spiritual needs of the elderly (Well-Spring Retirement Community Employee Manual, 2014). The main goals of the organization are to provide high and consistent quality services to its residents and maintain a safe and positive work environment for its employees. To accomplish its goals, Well-Spring has launched a series of annual WWPs to inform employees about healthy lifestyles and motivate them to engage in healthier behavior.

This study evaluates the impact of Well-Spring's "Just10" program that took place from February 3, 2014 through March 17, 2014. All employees were eligible to participate, but participation and completion were voluntary. The program taught employees how small commitments could yield large payoffs, such as an improved energy level and physical flexibility and mobility. The Just10 program was structured as a series of three challenges: losing ten pounds (lbs.), exercising ten minutes every day, and walking or running ten miles every week. Participants in the challenges were required to fill out activity logs provided by the Aquatic and Fitness Center Coordinator. Participants received $100 if they completed one or two challenges and received $300 if

they completed all three challenges. To record participant progress, the lose ten lbs. group was weighed before and after the program by the coordinator. Also, the exercise groups had to turn in activity logs that recorded the distance or the time of exercise to calculate how far they had walked or run and how long they had exercised. Employees with the highest percentage of weight loss in the lose ten lbs. group, and the most days of exercise and the most miles completed in the exercise groups won the cash rewards.

The Logic Model

In the program evaluation literature, a logic model is used to connect program resources to activities performed and outcomes achieved (Wholey, 1983; Rush and Ogborne, 1991). The aim is to identify components of the program and analyze how these components are associated. The logic model presented in Figure 4.1 illustrates the hypothesized causal relationship between Just10 program participation and the long-term health outcomes associated with the program.

The inputs are the resources that are required to implement and operate the program. Inputs of the Just10 program included labor and non-labor costs, which were measured from the company's perspective. The labor cost included the fitness center coordinator's time spent on program preparation, employee training, monitoring and recording participants' progress and results, as well as the maintenance and repair cost of the gym at the work site. The non-labor cost included incentives, which were cash rewards provided to motivate individuals to participate, and the space used for employee trainings. We provide a detailed explanation of program costs and their dollar values in the results section.

The intervention component of the logic model shows the information provided and the activities done during the Just10 program. Participants were encouraged to improve their physical health by consuming fewer daily calories, eating healthier foods, refraining from eating at least three hours before bedtime, and implementing planned and structured exercises. The program's inputs, combined with intervention activities, were intended to raise employees' awareness of how to lose weight and how to exercise with the purpose of creating a healthier workplace environment.

The output component of the model is the direct consequence of the evaluated program. There are three main program outputs identified in this model: weight loss, regular exercise, and increased physical activity. The activities related to changing one's diet were expected to result in weight loss. The activities related to physical activity were expected to result in more frequent and regular exercise in addition to weight loss.

Proximal outcomes are the expected changes in the health risks and behavior caused by the program. The proximal outcomes are caused by either program outputs or intervention information and activities, even if the program outputs are not achieved. Because of these channels, proximal outcomes of decreased risk to physical health and increased awareness of heathier lifestyles were expected. Although the proximal outcome measures are important in the model, they are unobserved in this study because the data are not available. The employee biometrics data were available to the health insurance provider. However, we could not obtain the data due to the Health Insurance and Privacy Accountability Act (HIPAA) of 1996, which regulates the security standards for

protecting health information. Thus, even though the logic model contains the proximal outcomes, we do not analyze them when evaluating the program.

Distal outcomes are the long-term desired impacts of the wellness program. The distal outcomes are caused by either the proximal outcomes, or outputs even if changes in health risk or behavior are not achieved. In this study, the distal outcomes, which are the analysis outcomes, included decreased turnover rate, decreased number of work-related injuries, decreased lost labor days, and an aggregate outcome of decreased organizational cost through decreased turnover and injuries.

To strengthen causal inference, the logic model needs to include external components such as confounders and moderators that could affect the causal relationship. The confounder factors are associated with both outcomes of interest and program participation. The confounders could affect the strength of association between Just10 and the distal outcome. The moderator factors (interaction terms with program participation) impact the strength or the direction of the causal relationship between program participation and hypothesized outcomes. For this study, the confounder and moderator factors were age, gender, fulltime status, years of experience at Well-Spring, and department (health care or independent living). Older females among participants were possibly at a disadvantage when it came to losing weight for natural reasons such as hormones and metabolism (Sattler et al., 2018). Moreover, the older the person's age, the lower his or her muscle mass, which might have been an obstacle to losing weight or exercising more for older participants (Institute of Medicine, 2003). The fulltime status and years of experience were expected to impact employee engagement in the program.

Fulltime and more experienced employees may be more likely to invest in wellness programs at work. The employee's department may be related to the intensity of the employee's job. For example, the health care department is subject to physically heavier workloads when compared to the independent living department, which could have impacted participation negatively. We provide detailed information on confounders, moderators, and outcomes in the data section.

Data

The Well-Spring administration provided individual-level, de-identified panel data on all employees for years 2013-2014. This study was reviewed and approved by the UNCG IRB. We used SAS statistical software (SAS version 9.4, Inc., Cary, NC) for the descriptive statistics and Stata Statistical Software (Release 14) for the outcome analysis. The dataset includes information on gender, birth year, fulltime status, department, dates of hire and termination, program participation and completion, dates of injury, injury-related modified duty days (present at work but cannot perform normal job requirements), injury-related lost work days (not present at work), hourly wages, hours of work per day, and overall cost of turnover based on the position. Date of injury, modified duty days, lost work days, hourly wage, and hours of work per day were only available for injured employees. Age, department indicators (health care and independent living), turnover rate, injury rate, lost labor days and organizational cost are derived using the dataset.

To analyze the impacts of the program, turnover, injury, lost labor days, and organizational costs are used as analysis outcomes. Turnover is an indicator of employees

being terminated in a period. Cost of turnover includes fees for drug screens, background checks, registry checks, time paid for orientation, training, initial shadowing days on units, advertising, and lag time while positions are open. Injury is an indicator of employees being injured at work in a period. Lost labor days are the sum of modified duty days and lost work days that occurred due to injuries. Cost of injuries are calculated using hourly wage, hours of work per day, and lost labor days due to any injury. Organizational costs are calculated as the sum of injury and turnover costs.

Periods are defined as follows: The "pre-program period" is between 2013 and the beginning of the Just10 program in February 3, 2014. The "program period" starts with the Just10 program and ends with the beginning of the 2015 wellness program, Triple Challenge, in February 2, 2015. Although the program period for evaluation was 12 months, Just10 took place from February 3, 2014 through March 17, 2014. Periods were determined based on the start times of wellness programs to separate the outcomes impacted from one or two wellness programs.

The study sample consists of 858 employee-periods. There are 116 participants in the Just10 program and 323 non-participants who were employed during the Just10 program. In the pre-program period, 303 out of 323 non-participants and all participants were employed at Well-Spring. We did not include the period that starts with the Triple Challenge wellness program, because the outcomes would be affected by both Just10 and Triple Challenge programs. The timeline for the evaluation and the program periods are presented in Figure 4.2.

The descriptive statistics in Table 4.1 compare participants with non-participants for both pre-program and program periods. There were significantly more females and fulltime workers in the participants compared to non-participants for both periods. Additionally, program participants were significantly older than the non-participants for both periods. The mean age was 43 years old in the pre-program period and 44 years old in the program period for participants. The mean age was around 36 years old in both periods for non-participants. Participants were significantly older compared to nonparticipants in both periods. Around two-thirds of the employees were in the health care department, and around one-third were in the independent living department. Groups were not significantly different based on departments. Participants were significantly more experienced compared to nonparticipants.

Table 4.1 also presents the descriptive statistics for analysis outcomes. Turnover rate is zero in the pre-program period due to how the study sample is defined. In the program period, turnover is significantly lower for the participant group. Injury rate, lost labor days, and organizational cost were lower for participants compared to non-participants for both periods.

Methods

Econometric Model

Program (Treatment) Effect

The economic evaluation literature of workplace wellness programs has commonly used the difference-in-differences (DID) method to deal with selection. The method eliminates individual-specific fixed effects and common time trends. The

individual-specific fixed effects are eliminated by differencing outcomes across periods. The common time shocks are eliminated by differencing outcome across participant and non-participant groups.

In this analysis, we use the DID regression model to estimate a program effect for the outcomes of interest. Because the outcomes are nonlinear, identifying program causal effects on the outcome is different than linear models. The following subsections lay out the standard DID model, the DID model for nonlinear outcomes, and deriving and estimating program effects for all outcomes.

Standard Difference in Differences Model

A standard linear DID model for two-group and two-period set up is as follows (Athey and Imbens, 2006; Puhani, 2008):

$$Y_{it} = P_{it}G_{it}Y_{it}^{1} + (1 - P_{it}G_{it})Y_{it}^{0} \tag{4.1}$$

$$Y_{it}^{1} = \delta_i + \alpha_1 G_{it} + \alpha_2 P_{it} + X_{it}\beta + \varepsilon_{it} \tag{4.2}$$

$$Y_{it}^{0} = \alpha_1 G_{it} + \alpha_2 P_{it} + X_{it}\beta + \varepsilon_{it} \tag{4.3}$$

Y_{it} is the outcome for employee i observed at period t and values are independent across individuals. P_{it} is the indicator for program period. G_{it} is the indicator for participation group. $P_{it}G_{it}$ is the interaction term, and the indicator for participant group in the program period. Y_{it}^{1} is the outcome when the program is present. Y_{it}^{0} is the outcome when the program is absent. $X_{it}\beta$ is the vector of control variables. ε_{it}.is the error term.

Y_{it} is the realized outcome, whereas Y_{it}^1 and Y_{it}^0 are the outcomes for potential program status. To derive the realized outcome, plug (2) and (3) into (1):

$$Y_{it} = \delta_i P_{it} G_{it} + \alpha_1 G_{it} + \alpha_2 P_{it} + X_{it}\beta + \varepsilon_{it} \tag{4.4}$$

In the linear DID model, average program effect is identified if the following assumptions hold (Blundell and Dias, 2009; Lechner, 2011):

Assumption 1: $E(\varepsilon_i \mid G=1, P, X) = E(\varepsilon_i \mid G=0, P, X) = E(\varepsilon_i)$ and $E(\varepsilon_i) = 0$

Assumption 2: $E(\delta_i \mid G=1, P, X) = E(\delta_i \mid G=0, P, X) = E(\delta_i)$ and $E(\delta_i) = \delta$

Assumption 3: $E(Y_i \mid P=1, G, X) - E(Y_i \mid P=0, G, X)$

$$= E(Y_i \mid P=1, X) - E(Y_i \mid P=0, X)$$

The first assumption implies that the error term of the linear model is independent in the conditional mean. Because the error term is uncorrelated with observable and unobservable determinants of participation decision, there is no selection on the observables and the unobservables. The second assumption implies that the program effect (δ) is independent of participant group, so there is no selection into the program. The third assumption implies that if there was no program, both groups would have experienced the same time trends (common time trends assumption). The average program effect (average treatment effect - ATE), δ, is calculated as the cross differences of the outcome across periods and groups:

$$\delta^{ATE} = E(\hat{\delta}) = \left[E(Y \mid G=1, P=1, X) - E(Y \mid G=1, P=0, X)\right] - \left[E(Y \mid G=0, P=1, X) - E(Y \mid G=0, P=0, X)\right] \quad (4.5)$$

The program effect might be different across individuals. In that case, Assumption 2 is violated and $E(Y \mid G=1, P, X) \neq E(Y \mid G=0, P, X)$. Then, the standard DID estimator gives the average program effect for participants (Average treatment effect on treated - ATT) (Athey and Imbens, 2006).

Difference in Differences in Nonlinear Models

Outcomes of interests are nonlinear variables in the analysis. Thus, consider the following nonlinear form of equations (4.2) and (4.3):

$$Y_{it}^{1} = F(\delta_i P_{it} G_{it} + \alpha_1 G_{it} + \alpha_2 P_{it} + X_{it}\beta + \varepsilon_{it}) \quad (4.6)$$

$$Y_{it}^{0} = F(\alpha_1 G_{it} + \alpha_2 P_{it} + X_{it}\beta + \varepsilon_{it}) \quad (4.7)$$

where Y_{it}^{1} is nonlinear outcome variable under the program, and F(.) is a nonlinear transformation function. Y_{it}^{0} is the unobserved counterfactual outcome but modeled as equation (4.7). This specification keeps the linear index structure of standard DID model, but it has a nonlinear link function, F(.) .The program effect cannot be calculated as the cross differences as in equation (4.5) because of nonlinearity. Instead the ATT is as follows (Athey and Imbens, 2006; Puhani, 2008):

$$E(Y^1 \mid G=1, P=1, X) - E(Y^0 \mid G=1, P=1, X) \tag{4.8}$$
$$= F(\delta + \alpha_1 + \alpha_2 + X_{it}\beta + \varepsilon_{it}) - F(\alpha_1 + \alpha_2 + X_{it}\beta + \varepsilon_{it})$$

Equation (4.8) shows that we can still focus on the coefficient of the interaction term as program effect in the nonlinear specification because ATT is zero if and only if δ is zero. The program effect is the incremental effect of the coefficient of the interaction term, δ.

For nonlinear outcomes, additional assumptions are needed for identification of the ATT. The outcome function is assumed to be strictly monotone in ε, given period *t* and X_{it}. Strict monotonicity means higher unobservables correspond to strictly higher outcome assuming the unobservables are related to better health and higher ability levels. In addition, outcomes do not directly depend on the program participation, where outcome of an observation in the absence of the program also satisfies the functional form F(.). Under these assumptions, the sign of δ is the sign of the program effect. However, the coefficients do not show the magnitudes of the program effect due to nonlinearity (Athey and Imbens, 2006; Puhani, 2008).

Empirical Model

We estimate the following nonlinear outcome equation:

$$Y_{it} = F(\delta period_{it} just10_{it} + \alpha_1 just10_{it} + \alpha_2 period_{it} + \beta X_{it} + \varepsilon_{it}) \tag{4.9}$$

In the nonlinear form, δ is not the cross differences as in the linear DID model. Although the coefficient of the interaction term is not the program effect itself in the

nonlinear case, we can still use interaction term for interpretations because the average program effect on participants only exists when $\delta \neq 0$ as shown in equation (4.8). We use the marginal effects to interpret the effect of the interaction term on the nonlinear outcome

Model Specification

When the groups are not randomized, participants are likely to be systematically different from non-participants with their (1) observed characteristics, such as age and gender, (2) unobserved characteristics, such as genetical factors and health history, and (3) expected benefits from the wellness program that drives motivation. The DID approach controls for individual-specific effects and common time trends. In addition, we controlled for various confounders and moderators listed in the logic model as age, gender, fulltime status, years of experience at Well-Spring, and health care department.

The first specification includes interaction term of program participation and program period indicators ($period_{it}\, just10_{it}$), Just10 participation ($just10_{it}$), and period indicators ($period_{it}$). This estimation deals with the selection due to individual and time-specific effects. The second specification includes confounders (X_{it}) as well as interaction term, Just10 participation, and period indicators. This estimation deals with selection due to observed characteristics as well as individual and time-specific effects. The third specification includes moderators ($period_{it}\, just10_{it}\, X_{it}$), and confounders as well as interaction term, Just10 participation, and period indicators. This estimation deals with selection due to observed characteristics as well as individual and time-specific

effects. Furthermore, moderators, which are the interaction terms of confounders with $period_{it}\, just10_{it}$, indicates the differences in effects for age, gender, fulltime status, health care department, and experience groups. A joint test on moderators was performed to see whether there were any significant differences in these groups.

We estimate equation (4.9) for the analysis outcomes of turnover, injury, lost labor days, and organizational costs using the specifications explained above. Estimation methods are explained in the following subsections.

Turnover

Turnover ($turnover_i$) is the binary outcome that takes on a value of one if employees are terminated, and zero otherwise. We estimate the outcome using the logit model. The estimators are driven using maximum likelihood estimator. The log likelihood function is

$$\ln L(\theta) = \sum_{i=1}^{N} turnover_i \ln p_i + (1 - turnover_i) \ln(1 - p_i) \tag{4.10}$$

$$p_i = \Pr(turnover_i = 1 \mid W_i) = \Lambda(W_i\theta) = \frac{\exp(W_i\theta)}{1 + \exp(W_i\theta)} \tag{4.11}$$

where p_i is the predicted probabilities and $\Lambda(W_i\theta)$ is the cumulative density function with $0 < \Lambda(W_i\theta) < 1$. W_i is the (row) vector of covariates $[period_{it}\, just10_{it}, period_{it}, just10_{it}, X_{it}]$ and θ is the corresponding (column) vector of

parameters $[\delta \quad \alpha_1 \quad \alpha_2 \quad \beta]'$. The first order conditions for the logit maximum likelihood estimator (MLE) is

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \sum_{i=1}^{N} (turnover_i - p_i) W_i = 0 \tag{4.12}$$

Details for the first order conditions are provided in the Appendix 4.A. The marginal effect for the j^{th} regressor is

$$\frac{\Delta p_i}{\Delta W_{ij}} = p_i(1 - p_i)\theta_j \tag{4.13}$$

Equation (4.13) shows the change in the conditional mean of turnover when the j^{th} regressor change by one unit.

Injury

Injury ($injury_i$) is the binary outcome that takes on a value of one if employees are injured, and zero otherwise. We estimate the outcome using the logit model with the same procedure as in turnover estimations. The estimators are driven using MLE. The log likelihood function is

$$\ln L(\theta) = \sum_{i=1}^{N} injury_i \ln p_i + (1 - injury_i) \ln(1 - p_i) \tag{4.14}$$

$$p_i = \Pr(injury_i = 1 | W_i) = \Lambda(W_i\theta) = \frac{\exp(W_i\theta)}{1 + \exp(W_i\theta)} \tag{4.15}$$

where p_i is the predicted probabilities and $\Lambda(W_i\theta)$ is the cumulative density function with $0 < \Lambda(W_i\theta) < 1$. W_i is the (row) vector of covariates $[period_{it}\ just10_{it}, period_{it}, just10_{it}, X_{it}]$ and θ is the corresponding (column) vector of parameters $[\delta \quad \alpha_1 \quad \alpha_2 \quad \beta]'$. The first order conditions for the logit maximum likelihood estimator (MLE) is

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \sum_{i=1}^{N} (injury_i - p_i) W_i = 0 \tag{4.16}$$

Details for the first order conditions are provided in the Appendix 4.A. The marginal effect for the j^{th} regressor is

$$\frac{\Delta p_i}{\Delta W_{ij}} = p_i(1 - p_i)\theta_j \tag{4.17}$$

Equation (4.17) shows the change in the conditional mean of injury when the j^{th} regressor change by one unit.

Lost Labor Days

Lost labor days ($lostdays_i$) is a count variable that takes non-negative values. We estimate the outcome using the Poisson model. The estimators are driven using MLE. The log likelihood function is

$$\ln L(\theta) = \sum_{i=1}^{N} [\mu_i - lostdays_i \ln(\mu_i) - \ln(lostdays_i !)] \tag{4.18}$$

where μ_i>0 is the expected value or average of $lostdays_i$. The average μ_i may depend on the values of the covariates, W_i. We follow a conventional approach in Poisson modeling and set $\mu_i = \exp(W_i\theta)$ (Woolridge, 2010). W_i is the (row) vector of covariates $[period_{it}\ just10_{it}, period_{it}, just10_{it}, X_{it}]$ and θ is the corresponding (column) vector of parameters $[\delta \quad \alpha_1 \quad \alpha_2 \quad \beta]'$. The first order conditions for the Poisson MLE is

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \sum_{i=1}^{N} (\exp(W_i\theta) - lostdays_i)W_i = 0 \tag{4.19}$$

Details for the likelihood function first order conditions are provided in the Appendix 4.A. The marginal effect for the j^{th} regressor is

$$\frac{\Delta \mu_i}{\Delta W_{ij}} = \theta_j \exp(W_i\theta) \tag{4.20}$$

Equation (4.19) shows the changes in the conditional mean of lost duty days when the regressor change by one unit.

Even though there are many zeros in the data for this variable, we do not need to model excess zeros because we know that lost labor days are positive only if there is an injury. Injury is the zero-inflation process for the count process of lost labor days. Thus, we estimate lost labor days outcome conditional on injury.

Organizational Costs

Organizational cost ($cost_i$) is a continuous variable that takes nonnegative values. Cost could be zero because there is no injury and no turnover. Cost could also be zero if there is no turnover but there is an injury without lost labor days due to injuries. Because most employees remain with Well-Spring and do not have an injury, the distribution of organizational costs has a point mass at zero. We estimate the outcome using two-part model that deals with the point mass at zero and permits zeros and non-zeros to be generated by different densities. The estimators are driven using MLE. The first part models the probability of a zero and estimated using the logit model. The second part models the outcome conditional on being positive and estimated using generalized linear model (GLM). The log-likelihood function is

$$\begin{aligned}\ln L(\theta) = \sum_{i=1}^{N} & I(cost_i = 0)\ln(1-\Lambda(W_i\theta)) \\ & + I(cost_i > 0)\ln\{\Lambda(W_i\theta)g(W_i\gamma)\}\end{aligned} \tag{4.21}$$

where *I(A)* is indicator function, equal to 1 if *A* is true, and 0 otherwise. The first part has the cumulative distribution function of logit model, $\Lambda(.)$, and the second part has gamma distribution with log link function, $g(.)$. Details for the likelihood function and first order conditions are provided in the Appendix 4.A. The marginal effect of the second part GLM estimations for the j^{th} regressor is (Frondel and Vance, 2012):

$$\frac{\Delta(\Lambda(W_i\theta)g(W_i\gamma))}{\Delta W_{ij}} = \theta_j \frac{\Delta(\Lambda(W_i\theta))}{\Delta W_{ij}} g(W_i\gamma) + \Lambda(W_i\theta)\frac{\Delta(g(W_i\gamma))}{\Delta W_{ij}} \tag{4.22}$$

We used a generalized linear model (GLM) with log link and gamma distribution to accommodate the heavy tail (positive skew) in the data. We chose GLM over ordinary least squares with log-scale because GLM offers unbiased estimates but might result in a loss of efficiency when errors are heteroscedastic (Belotti et al, 2015, Manning and Mullahy, 2001; Buntin and Zaslavsky, 2004).

Return on Investment

ROI was measured as the ratio of net benefit (the difference between benefit and cost) to cost, which had a threshold for positive ROI of zero. Savings from organizational costs were the benefit component and the program costs were the cost component. To calculate ROI, separate regressions of organizational costs and program costs were estimated.

Benefit was the marginal effect on the interaction term from the two-part model explained above. We are interested in overall ROI for the company. Thus, we only considered the confounders as control variables but not the moderators when we estimated the organizational cost for the ROI. Cost was the coefficient on the program participation repressor from ordinary least squares (OLS) regression. Program costs were regressed only in the program participation. In the benefit estimation, the negative sign of the marginal effect means the organization saved money from program participation in the program period. If the sign of the marginal effect is positive, it means the organization lost money from program participation in the program period. Therefore, when calculating ROI, we use $((-benefit - cost)/cost)$ formula.

ROI is formulated as a ratio; therefore, statistical precision information around ROI cannot be estimated. We used the nonparametric bootstrapping method to calculate confidence intervals around ROI. Bootstrapping provides ways of estimating confidence intervals and other measures of statistical precision by building data of replicated estimations. Organizational and program costs were estimated for each bootstrap replicate, resulting in an estimate of the ROI for each replicate. We regressed program costs instead of just using the average program cost because it was necessary to estimate ROI for each bootstrap replicate.

The nonparametric bootstrap method makes no assumption on distribution of the original data, which is advantageous when the sample size is small. Let $\hat{\delta}$ be the regressed value that is calculated using the original data. Let $\hat{\delta}_b$ be the values that is calculated from b^{th} bootstrap sample where $b = 1, 2, ..., B$ denotes the bootstrap sample. We used bias-corrected confidence intervals within the nonparametric bootstrapping process because it yields better result in case the statistics are biased. The bias-corrected 95% confidence interval is (Cameron and Trivedi, 2010; Hamilton, 1991)

$$CI : \left[\Phi(2z_0 - z_{0.025}), \Phi(2z_0 + z_{0.975})\right]$$

$$z_0 = \Phi^{-1}(\#\{\hat{\theta}_b \leq \hat{\theta}\})$$

where $\#\{\hat{\theta}_b \leq \hat{\theta}\}$ is the number of elements of the bootstrap sample distribution that are less than or equal to the observed sample distribution, and $\Phi(.)$ is the cumulative density distribution of standard normal variable.

Results

Program and Organizational Costs

Program costs, which are measured from the perspective of Well-Spring, are the inputs related to implementing the Just10 program. The units of measurement, unit costs and the total cost in 2015 dollars are presented in Table 4.2. Program costs are calculated using economic costing which includes opportunity costs of labor and space used during Just10 as well as the accounting cost of cash prizes. Labor costs include the fitness coordinator's hours and cost of maintenance and repair of the gym at the worksite. The fitness coordinator spent 4 hours per week (16 hours per month on average) working for the program. The hourly wage of the fitness coordinator was $17.52. Cost of the gym maintenance and repair was $11.42 per user per month. Total number of users was 123 including 102 residents and 21 employees. The total labor cost for the entire one and half months of the program was $780.21.

Non-labor costs contain the space used to train employees and prizes provided upon completion of the program. The fitness coordinator used the conference room at Well-Spring for the trainings. The opportunity cost of not renting the space is the fixed rental fee for a half-day, which was $395. Prizes were provided based on the number of completed challenges. Participants who completed less than three challenges won $100

and participants who completed all three challenges won $300. There were 40 participants who received $100 and 34 participants who received $300. Total non-labor cost of the program was $14,612. Overall, the program cost Well-Spring $15,392, or $132.69 per participant, to implement.

Organizational cost includes monetized values of employee and workplace outcomes improved by the Just10 program, such as costs of injury and turnover. Cost of workplace injuries are calculated by multiplying lost labor days, hours of work per day, and hourly wage. Terminated employees were assigned a turnover cost calculated from the midpoint of the following cost ranges based on their position: wait staff cost ranges from $0 to $500, in line staff cost ranges from $2,000 to $4,000, staff in professional or skilled positions cost ranges from $3,000 to $8,000, and some department managers cost ranges from $7,000 to $50,000. Wait stuff could include voluntary interns which resulted in zero turnover cost to Well-Spring when they completed their internship and left. The organizational costs are used in the ROI analysis and the program costs are used in the ROI analyses. All cost measures are adjusted to 2015 US dollars using the Consumer Price Index from Bureau of Labor Statistics.

Outcome Analysis

We employ logit models to estimate turnover and injury outcomes, a Poisson model to estimate the lost labor days, and two-part model to estimate the organizational cost with three same specifications of no control, control for confounders, and control for both confounders and moderators. All models are estimated with cluster-robust standard errors. Confounders include age, experience at Well-Spring, and indicators for female,

fulltime, and health care departments. Although the participant and nonparticipant groups are not systematically different by the department, we control for the department because we expect it to impact the outcomes as discussed earlier. Moderators include interactions of Just10 participation and program period with age, female, fulltime employment, health care department, and Well-Spring experience variables (three-way interaction terms). We tested whether the moderators that are included in the analysis are jointly significant.

Turnover

Table 4.3 presents coefficients for analysis sample and marginal effects for average program effect on participants from the regressions of turnover. Marginal effects were calculated as discrete change from the base level. Recall that there is no turnover in the pre-program, by definition. Therefore, when we estimated the DID specification, only program participation was included. Just10 participation in the program period significantly decreases the probabilities of turnover around 9 percentage point in specification (1) but has the opposite effect in specifications (2) and (3). The coefficient and marginal effect for the interaction effect are not significant except for the first model.

Employees with fulltime status or more experience are less likely to be terminated. If an employee has fulltime status, turnover significantly decreases by 11-15 percentage points at 1% level compared to a part time employee. If an employee has one or more years of experience at Well-Spring, turnover significantly decreases around 5-6percentage points at 1% level. Moreover, if an employee works in the health care department, turnover significantly decreases around 5-9 percentage points at 10% level.

Even if moderators are jointly significant, they have no significant impact on the outcome.

Injury

Table 4.4 presents coefficients for analysis sample and marginal effects for average program effect on participants from the regressions of injury. Marginal effects were calculated as discrete change from the base level. Just10 participation in the program period significantly decreases the probabilities of injury by 2 percentage points in specifications (1) and (2), and 3.5 percentage points in specification (3). Because injury is a rare event, the standard errors are quite high, and the effect of the interaction term is not statistically significant.

Specification (3) shows that moderators are jointly significant and have significant effect on the outcome at 1% level. The presence of significant moderators in column (3) indicated that variables are now conditional on values of the variables they are interacted.

Injury is a rare event, which leads to underestimated probabilities and maximum likelihood estimates suffer from small-sample bias (King and Zeng, 2001). Although the coefficient of Just10 participation in the program period changes due to three-way interactions, listed above, high magnitude in the coefficient seems questionable. Injury being rare event and three-way interaction terms together contribute to a large bias for the coefficient of Just10 participation in the program period.

Lost Labor Days

Table 4.5 presents coefficients and marginal effects for the analysis sample from the regressions of lost labor days. Marginal effects were calculated as discrete change from the base level. The regressions were run conditional on having an injury. Thus, the sample size in the regressions were very small which greatly impacted the statistical significance. Participating in Just10 in the program period decreases the number lost labor days in all specifications around 55 days. In specification (3), the moderators are omitted due to collinearity. The marginal effect of participating in Just10 in the program period is not estimable.

When an injury occurs, lost labor days significantly increase as age goes up. Similarly, employees with fulltime status has significantly higher lost labor days when an injury occurs. Employees with more experience have less lost labor days after an injury. Most of the coefficients were significant because we only estimate the outcome using the injured sample (22 observations). Marginal effects were significantly high in magnitude due to the same reason.

The marginal effect on Just10 participation in the program period is high in magnitude. This is due to lost labor days is being rare event and having extreme outliers. We ran the regression on only injured sample, which included 22 observations with mean 54 days lost, minimum 0 days lost, and maximum 554 days lost due to injuries. Although, estimates might suffer from inflated effects, bias in maximum likelihood estimations, and low statistical power, the marginal effects are actually close to sample mean.

Organizational Costs

Tables 4.6 and 4.7 present coefficients from the first and second part GLM regressions for analysis sample and marginal effects for average program effect on participants from the organizational cost regressions. The method requires retransformation of log organizational costs into organizational costs. Thus, we used marginal effects when we interpreted the significance and the magnitudes for the right-hand side variables. Marginal effects were calculated as discrete change from the base level. The logit estimations show that participating in Just10 in the program period decreases the probability of having positive organizational costs. The second part shows that, conditional on having organizational costs, participating in Just10 in the program period slightly increased the organizational costs. After the retransformation, marginal effects show that participation in the program period decreased the organizational costs in specification (2) and (3) but increased in specification (1). The program effect was not significant in any specification.

Organization costs per employee goes up significantly at 10% level as age of employee increases. Female employees and employees with higher experience have significantly lower organizational costs. Employees in health care department have significantly higher organizational costs.

ROI Estimation

To calculate ROI, we estimated the regressions of organizational and program costs. We did not apply any discounting because all costs accrued within the year so there is no need to account for the changing value of the dollar over time. To estimate

organizational cost, we used a two-part model where the first part is a logit model on probability of organizational costs being positive and the second part is GLM with log link and gamma distribution. The regressions are provided in tables 4.6 and 4.7. To estimate the average program cost, we used the OLS model with only Just10 participation. The estimations are presented in Table 4.8.

ROI is measured as the difference between organizational costs and program costs, divided by the effect on program cost. Details on ROI calculation is provided in the methods section. The 95% confidence interval, which was the 2.5 and 97.5 percentiles of the distribution around the organizational costs, program costs, and ROI were estimated using a nonparametric bootstrapping with 1000 replications. The estimated organization cost is -$311.004 which is the program benefits with a confidence interval of [-1595.915, 11636.34]. The estimated program cost is $132.69 with a confidence interval of [112.957, 156.101]. Then the ROI is [(-(-$311.004)-$132.69)/ $132.69] = 1.344with the bias-corrected confidence interval [-13.271, 84.145]. The wide range of confidence intervals in ROI is due to high variability in organizational cost. Although Well-Spring saved $2.344 for every $1 dollar of program investment, the wide range of confidence intervals for benefits, and therefore ROI reduce the precision of the ROI.

Sensitivity Analysis for ROI Estimates

We calculated the organizational cost using the sum of turnover and injury costs. We used midpoint of turnover cost ranges. In the sensitivity analysis, minimum and maximum point of turnover cost ranges were included in the organizational costs, respectively. Tables 4.9 and 4.10 presents the results. When organizational costs include

minimum of turnover cost, then ROI is (-$1.836) with the bias-corrected confidence interval [-5.960, 32.253] and it means Well-Spring lost $0.836 for every $1 dollar of program investment. When organizational costs include maximum of turnover cost, ROI is (-$0.989) with the bias-corrected confidence interval [-17.425, 84.844] and it means Well-Spring saved only $0.011 for every $1 dollar of program investment. The wide range of confidence intervals for benefits, and therefore ROI reduce the precision of the ROI in sensitivity analysis as well.

ROI estimations are sensitive to turnover costs because the difference between minimum and maximum turnover cost increases as the position of the employee ranks higher. For example, this difference is $500 for the wait stuff position, where as it is $43000 for the department manager position.

Discussion

The economic evaluation literature of WWP has been criticized because of its lack of significant cost savings or positive ROI and issues with the choice of analytic methods to estimate outcomes. In addition, ROI of workplace wellness programs were usually evaluated in large companies which does not represent an average organization size. The systematic review chapter of this dissertation showed that there were only four studies that reported confidence intervals around ROI for its significance, seven studies that used proper analytic method based on outcome distributional properties, and three studies with ROI of workplace wellness programs in small companies. The literature needs independent research to fill this knowledge gap, which negatively impacts adopting WWPs, and provide reliable evidence.

The evaluation of WWP at Well-Spring presented the ROI of a multicomponent wellness program in a small nonprofit company. The main analysis accounted for differences in turnover, injury, lost labor days, and organizational costs. Turnover outcome was monetized using cost ranges provided by Well-Spring. Injury outcome was monetized using lost labor days, hours of work per day and hourly wages. We estimated the organizational costs savings of $210.342 per participant over a two-year period. The overall ROI was 0.585 (95% confidence interval, (-35.095, 14.103) or 58.5%, indicating that average organizational costs declined by $1.585 for every $1.00 spent on the Just10 wellness program. The systematic review analysis showed that included articles had an average ROI of 0.67 for small companies (Maniscalco et al., 1999; Palumbo et al., 2013; Griffin et al., 2016). Although the ROI of Well-Spring was below the ROI average of systematic review, Well-Spring did indeed save money from the wellness program. However, the large confidence interval around the ROI estimate, estimated using a bootstrapping method, showed that the results are not precise.

Five main limitations need to be pointed out. First, the program participation in exercising groups was self-reported, which might be overstated by participants. Completions of exercising challenges were rewarded with cash prizes. Program costs might have gone up due to measurement error. Measurement errors in the explanatory variable due to self-reporting could lead to bias estimates (Gujarati, 2009). Second, the ROI finding may have been understated or overstated due to missing compensation claims data. At the time the data were provided, the compensation claims were not available to the company. Third, cost of turnover only varies by the employee position

not by individual-level which would impact the organizational cost when calculating ROI. Measurement errors in the dependent variables due to data observability issues produce unbiased estimates; however, the power of statistical tests is reduced due to larger variances (Gujarati, 2009). Fourth, despite the effort to reduce the selection bias using delivering DID specification, the estimation results may be biased due to heterogenous program effects. Fifth, injury and lost labor days outcomes are rare events which could also impact the significance and magnitude of the findings. Rare events lead to bias estimates (King and Zeng, 2001). Lost labor days and organizational costs had extreme outliers which impact the efficiency of the estimates due to heteroscedasticity. In a rare event case, the effect of extreme outliers on estimations could magnify.

The systematic review chapter shows that the statistical inference information for ROI are mostly missing and large companies are overrepresented in the literature. In addition, articles on small companies had an average quality score of 7.4 where the overall average score of included articles were 11.5 out of 18. Well-Spring study contributes to the field by providing rigorous evidence on a small nonprofit company and presenting the confidence intervals around ROI estimate. This essay scores 15.5 out of 18 when evaluated based on the quality rigor rubric provided in Chapter II. Although the observational design negatively impacted the score, this paper scored above the average of randomized studies, which was 14.2, included in the systematic review. Therefore, we can argue that the field can improve the study rigor even if more randomized study cannot be produced. This study can be generalized to a larger population of small

companies in terms of program evaluation methods. However, findings do not necessarily imply to WWPS in other small companies.

The WELCOA chapter of this dissertation shows that there are misalignments in the economic evaluation literature of WWPs in terms of companies' reasons for adoption and evaluated outcomes. Well-Spring study contributes to the field by providing rigorous evidence on a small nonprofit company, by aligning the objectives and evaluated outcomes, and by presenting the confidence intervals around ROI estimate. The contributions are expected to shift the understanding of why and how we should evaluate WWPs and to assist employers deciding on provisions of WWPs.

Figures

Figure 4.1. Logic Model for the Just10 Program

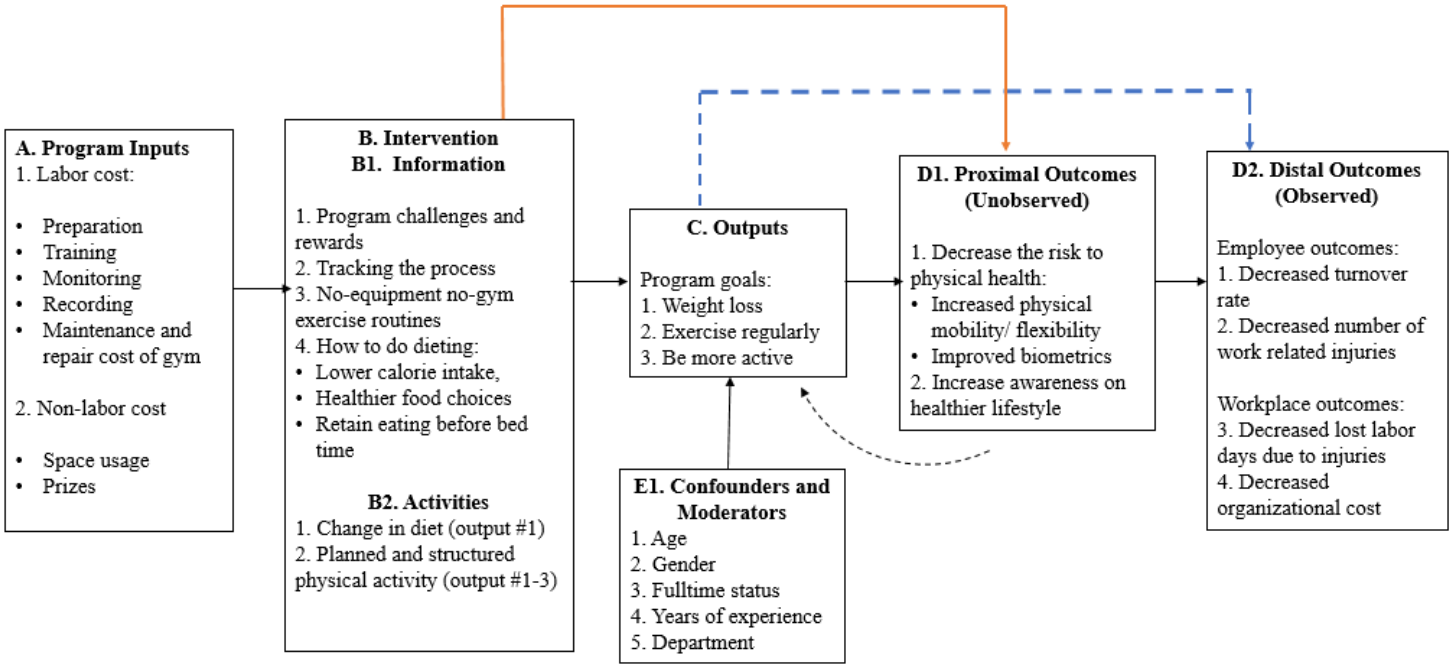
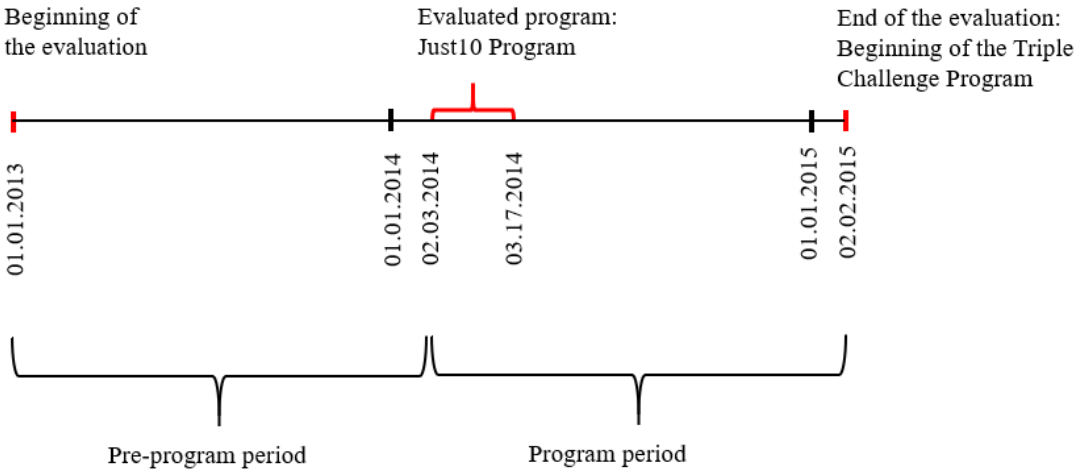


Figure 4.2. Timeline for the Evaluation and the Programs



Tables

Table 4.1. Descriptive Statistics of Study Sample by Period and Just10 Participation

	Pre-program period			Program period		
	PJ10	NPJ10	Difference	PJ10	NPJ10	Difference
Age	42.862	35.716	-7.146***	43.862	36.344	-7.518***
	(12.90)	(16.06)	(1.666)	(12.90)	(15.92)	(1.643)
Female	0.897	0.762	-0.134***	0.897	0.771	-0.126***
	(0.306)	(0.426)	(0.043)	(0.306)	(0.421)	(0.043)
Fulltime	0.793	0.426	-0.367***	0.793	0.415	-0.378***
	(0.407)	(0.495)	(0.052)	(0.407)	(0.493)	(0.051)
Independent living	0.319	0.360	0.041	0.319	0.347	0.028
	(0.468)	(0.481)	(0.052)	(0.468)	(0.477)	(0.051)
Health care	0.681	0.640	-0.041	0.681	0.653	-0.028
	(0.468)	(0.481)	(0.052)	(0.468)	(0.477)	(0.051)
Well-Spring experience	8.899	6.373	-2.526***	8.899	6.084	-2.815***
(years)	(6.693)	(5.433)	(0.634)	(6.693)	(5.390)	(0.624)
Turnover rate	0.000	0.000	0.000	0.172	0.263	0.091**
	(0)	(0)	(0.000)	(0.379)	(0.441)	(0.046)
Injury rate	0.034	0.030	-0.005	0.009	0.025	0.016
	(0.183)	(0.170)	(0.019)	(0.0928)	(0.156)	(0.015)
Lost labor days	0.948	2.271	1.322	0.000	1.223	1.223
	(6.416)	(32.21)	(3.015)	(0)	(14.59)	(1.355)
Organizational cost	78.557	164.546	85.988	838.793	1150.236	311.443
(2015$)	(528.2)	(2287.1)	(214.654)	(1889.8)	(2718.8)	(273.547)
Number of employees	116	303		116	323	

Notes: Mean coefficients; standard deviations are in parentheses. * p<0.10, ** p<0.05, *** p<0.01. Difference is calculated as the difference between averages for Just10 nonparticipants (NPJ10) and participants (PJ10).

Table 4.2. Cost of the Just10 Program

Measure	**Unit**	**Per Unit Price (2015 $)**	**No. of Units**	**Total (2015 $)**
1. Labor cost				
1.1. Personnel cost for preparation, training, monitoring, recording (monthly average)	per hour	17.52	16	280.32
1.2. Cost of gym maintenance and equipment (monthly average)	per employee	11.42	21	239.82
Total labor cost	per month	520.14	1.5	780.21
2. Non-labor cost				
2.1. Prizes	employee (less than three challenges)	100.12	40	4,004.80
	employee (three challenges)	300.36	34	10,212.24
2.2. Space (Well-Spring conference room)	half day	395.02	1	395.02
Total non-labor cost				14,612.06
Total program cost				15,392.27
Program cost per-participation	participants		116	132.69

Table 4.3. Estimation Results for Turnover

	(1) Logit	(2) Logit	(3) Logit
Just10 participation	-0.539*	0.503	0.060
	(0.277)	(0.349)	(1.866)
	[-0.091]**	[0.047]	[0.007]
Age		-0.001	0.009
		(0.011)	(0.012)
		[-0.000]	[0.001]
Female		0.192	0.212
		(0.364)	(0.409)
		[0.019]	[0.023]
Fulltime		-0.984***	-1.188***
		(0.321)	(0.398)
		[-	[-0.145]***
Health care department		-0.504*	-0.756**
		(0.265)	(0.300)
		[-0.054]*	[-0.088]**
Well-Spring experience		-0.472***	-0.556***
		(0.143)	(0.144)
		[-	[-0.062]***
Age*Interaction term			-0.040
			(0.027)
			[-0.004]
Female*Interaction term			0.366
			(1.022)
			[0.038]
Fulltime*Interaction term			0.973
			(0.692)
			[0.095]
Health care department*Interaction term			0.934
			(0.749)
			[0.095]
Experience*Interaction term			0.267
			(0.261)
			[0.030]
Constant	-1.030***	-2.730***	-2.903***
	(0.127)	(0.861)	(0.943)
Log pseudolikelihood	-239.480	-170.485	-167.276
Pseudo R-square	0.008	0.294	0.307
Number of employee-period	439	439	439
Number of employees	439	439	439

Notes: Cluster-robust standard errors are in parenthesis. Marginal effects for participant in the program period, which is the average treatment effect on treated, are in brackets. The sample for marginal effects is 116 participants. * p<0.10, ** p<0.05, *** p<0.001. Variable of interest is the program effect which is the Just10 participation. In pre-program period, turnover is zero because of the way the analysis sample was defined. Thus, interaction term and program period indicators are omitted. Age and Well-Spring experience are demeaned continuous variables. Periods are defined as follows: The pre-program period is between 2013 and the beginning of the Just10 program. The program period starts with the Just10 program and ends with the beginning of the 2015 wellness program, Triple Challenge

Table 4.4. Estimation Results for Injuries

	(1) Logit	(2) Logit	(3) Logit
Just10 participation*Program period (Interaction term)	-1.226 (1.230) [-0.020]	-1.230 (1.228) [-0.020]	-140.008*** (7.056) [-0.353]***
Just10 participation	0.154 (0.611) [0.001]	-0.209 (0.695) [-0.002]	-0.212 (0.695) [-0.000]
Program period	-0.187 (0.493) [-0.002]	-0.177 (0.499) [-0.002]	-0.195 (0.510) [-0.000]
Age		-0.011 (0.018) [-0.000]	-0.003 (0.016) [-0.000]
Female		0.347 (0.557) [0.003]	0.185 (0.534) [0.000]
Fulltime		1.355** (0.555) [0.008]	1.420*** (0.545) [0.000]
Health care department		-0.067 (0.447) [-0.001]	0.090 (0.444) [0.000]
Well-Spring experience		-0.026 (0.049) [-0.000]	-0.069* (0.040) [-0.000]
Age*Interaction term			-1.483*** (0.069) [-0.000]
Female*Interaction term			11.647*** (1.364) [0.001]
Fulltime*Interaction term			-18.396*** (1.917) [-0.029]**
Experience*Interaction term			10.521*** (0.493) [0.000]***
Constant	-3.486*** (0.339)	-4.554*** (0.658)	-4.628*** (0.669)
Log pseudolikelihood	-101.148	-97.828	-91.594
Pseudo R-square	0.011	0.044	0.105
Number of employee-period	858	858	858
Number of employees	439	439	439

Notes: Cluster-robust standard errors are in parenthesis. Marginal effects for participant in the program period, which is the average treatment effect on treated, are in brackets. The sample for marginal effects is 116 participants. * p<0.10, ** p<0.05, *** p<0.001. Variable of interest is the program effect which is the interaction term. Age and Well-Spring experience are demeaned continuous variables. Periods are defined as follows: The pre-program period is between 2013 and the beginning of the Just10 program. The program period starts with the Just10 program and ends with the beginning of the 2015 wellness program, Triple Challenge.

Table 4.5. Estimation Results for Lost Labor Days

	(1) Poisson	(2) Poisson	(3) Poisson
Just10 participation*Program period (Interaction term)	-15.082*** (1.455) [-55.033]**	-6.472 (4.336) [-54.143]***	-6.472 (4.336) N/A
Just10 participation	-1.022 (0.859) [-40.419]	-1.682** (0.847) [-61.966]**	-1.682** (0.847) [-61.966]**
Program period	-0.437 (0.955) [-22.689]	0.211 (0.569) [11.914]	0.211 (0.569) [11.914]
Age		0.089*** (0.026) [4.846]**	0.089*** (0.026) [4.846]**
Female		-3.208** (1.325) [-1003.374]	-3.208** (1.325) [-1003.374]
Fulltime		1.426 (1.134) [51.474]*	1.426 (1.134) [51.474]*
Health care department		2.418** (1.001) [219.034]	2.418** (1.001) [219.034]
Well-Spring experience		-0.400** (0.172) [-21.696]*	-0.400** (0.172) [-21.696]*
Age*Interaction term			0.000 (.) [0.000]
Female*Interaction term			0.000 (.) [0.000]
Fulltime*Interaction term			0.000 (.) [0.000]
Experience*Interaction term			0.000 (.) [0.000]
Constant	4.337*** (0.762)	2.828* (1.453)	2.828* (1.453)
Log pseudolikelihood	-1501.993	-773.777	-773.777
Number of employee-period	22	22	22
Number of employees	22	22	22

Notes: Cluster-robust standard errors are in parenthesis. Marginal effects are in brackets. * p<0.10, ** p<0.05, *** p<0.001. Variable of interest is the program effect which is the interaction term. Age and Well-Spring experience are demeaned continuous variables. Periods are defined as follows: The pre-program period is between 2013 and the beginning of the Just10 program. The program period starts with the Just10 program and ends with the beginning of the 2015 wellness program, Triple Challenge.

Table 4.6. The First Part Estimation Results for Organizational Cost

	(1) Logit	(2) Logit	(3) Logit
Just10 participation*Program period (Interaction term)	-1.203* (0.711)	-1.199* (0.728)	-2.407 (1.835)
Just10 participation	0.570 (0.655)	1.299* (0.720)	1.308* (0.734)
Program period	2.966*** (0.431)	3.171*** (0.467)	3.167*** (0.473)
Age		-0.005 (0.009)	0.000 (0.010)
Female		0.078 (0.301)	0.075 (0.321)
Fulltime		-0.575** (0.282)	-0.673** (0.312)
Health care department		-0.261 (0.227)	-0.371 (0.239)
Well-Spring experience		-0.299*** (0.077)	-0.299*** (0.075)
Age*Interaction term			-0.031 (0.026)
Female*Interaction term			0.502 (0.995)
Fulltime*Interaction term			0.458 (0.641)
Health care department*Interaction term			0.548 (0.724)
Experience*Interaction term			0.009 (0.231)
Constant	-3.902*** (0.413)	-5.069*** (0.725)	-4.910*** (0.771)
Log pseudolikelihood	-1424.739	-1347.113	-1340.799
Pseudo R-square	858	858	858
Number of employee-period	439	439	439
Number of employees	-1.203*	-1.199*	-2.407

Notes: Cluster-robust standard errors are in parenthesis. Marginal effects are in brackets. * p<0.10, ** p<0.05, *** p<0.001. Variable of interest is the program effect which is the interaction term. Age and Well-Spring experience are demeaned continuous variables. Periods are defined as follows: The pre-program period is between 2013 and the beginning of the Just10 program. The program period starts with the Just10 program and ends with the beginning of the 2015 wellness program, Triple Challenge.

Table 4.7. The Second Part Estimation Results for Organizational Cost

	(1) GLM	(2) GLM	(3) GLM
Just10 participation*Program period (Interaction term)	1.469* (0.796) [380.452]	0.570 (0.852) [-311.004]	0.603 (0.865) [-510.027]
Just10 participation	-1.294 (0.789) [-1031.915]	-0.644 (0.776) [491.313]	-0.440 (0.823) [342.895]
Program period	-0.711 (0.697) [733.675]***	0.609 (0.652) [1361.528]***	0.833 (0.654) [811.374]***
Age		0.022*** (0.007) [25.164]	0.023** (0.009) [19.571]**
Female		-1.086*** (0.343) [-2303.651]	-1.414*** (0.371) [-2382.912]*
Fulltime		0.913*** (0.301) [625.217]	1.243*** (0.441) [654.841]
Health care department		1.814*** (0.279) [1565.424]***	2.152*** (0.278) [1537.286]**
Well-Spring experience		0.020 (0.030) [-283.787]***	-0.014 (0.049) [-171.201]**

Notes: Cluster-robust standard errors are in parenthesis. Marginal effects for participant in the program period, which is the average treatment effect on treated, are in brackets. The sample for marginal effects is 116 participants. * p<0.10, ** p<0.05, *** p<0.001. Variable of interest is the program effect which is the interaction term. Age and Well-Spring experience are demeaned continuous variables. Periods are defined as follows: The pre-program period is between 2013 and the beginning of the Just10 program. The program period starts with the Just10 program and ends with the beginning of the 2015 wellness program, Triple Challenge.

Table 4.7. The Second Part Estimation Results for Organizational Cost (Continued)

	(1) GLM	(2) GLM	(3) GLM
Age*Interaction term			-0.030*** (0.010) [-41.647]**
Female*Interaction term			1.805*** (0.396) [923.209]***
Fulltime*Interaction term			-1.194*** (0.442) [-962.530]
Health care department*Interaction term			-1.944*** (0.318) [-2685.313]
Experience*Interaction term			0.031 (0.049) [31.298]
Constant	9.025*** (0.690)	7.385*** (0.704)	6.969*** (0.758)
Log pseudolikelihood	-292.246	-239.501	-238.278
Pseudo R-square	0.163	0.314	0.317
Number of employee-period	121	121	121
Number of employees	117	117	117

Notes: Cluster-robust standard errors are in parenthesis. Marginal effects for participant in the program period, which is the average treatment effect on treated, are in brackets. The sample for marginal effects is 116 participants. * p<0.10, ** p<0.05, *** p<0.001. Variable of interest is the program effect which is the interaction term. Age and Well-Spring experience are demeaned continuous variables. Periods are defined as follows: The pre-program period is between 2013 and the beginning of the Just10 program. The program period starts with the Just10 program and ends with the beginning of the 2015 wellness program, Triple Challenge.

Table 4.8. ROI Estimation

	Organizational cost (Benefit)	Program Cost (Cost)	ROI
Marginal cost estimate (per-employee)	-311.004	132.692***	1.344
Bias-corrected confidence intervals	[-1595.915, 11636.34]	[112.957, 156.101]	[-13.271, 84.145]
Number of employee-period	858	858	858
Number of employees	439	439	439

Notes: ROI = (-Benefit-Cost)/Cost. The negative sign shows that there were savings. Thus, when we calculated the ROI, we multiplied the coefficient with (-1).

Table 4.9. Sensitivity Analysis: Minimum Cost of Turnover

	Organizational cost (Benefit)	Program Cost (Cost)	ROI
Marginal cost estimate (per-employee)	-110.902	132.692***	-1.836
Bias-corrected confidence intervals	[-622.486, 4491.694]	[112.957, 156.101]	[-5.960, 32.253]
Number of employee-period	858	858	858
Number of employees	439	439	439

Notes: ROI = (-Benefit-Cost)/Cost. The negative sign shows that there were savings. Thus, when we calculated the ROI, we multiplied the coefficient with (-1).

Table 4.10. Sensitivity Analysis: Maximum Cost of Turnover

	Organizational cost (Benefit)	Program Cost (Cost)	ROI
Marginal cost estimate (per-employee)	1.623	132.692***	-0.989
Bias-corrected confidence intervals	[-2151.462, 11295.52]	[112.957, 156.101]	[-17.425,84.844]
Number of employee-period	858	858	858
Number of employees	439	439	439

Notes: ROI = (-Benefit-Cost)/Cost. The negative sign shows that there were savings. Thus, when we calculated the ROI, we multiplied the coefficient with (-1).

References

1. Adams, M. E., McCall, N. T., Gray, D. T., Orza, M. J., & Chalmers, T. C. (1992). Economic analysis in randomized control trials. *Medical Care*, 231-243.

2. Aldana, S. G. (2001). Financial impact of health promotion programs: A comprehensive review of the literature. *American Journal of Health Promotion,* 15(5): 296–320.

3. Anderko L, Roffenbender JS, Goetzel RZ, Millard F, Wildenhaus K, et al. (2012). Promoting Prevention Through the Affordable Care Act: Workplace Wellness. *Prev Chronic Dis*, 9.

4. Athey, S., & Imbens, G. W. (2006). Identification and inference in nonlinear difference-in-differences models. *Econometrica*, *74*(2), 431-497.

5. Baicker, K., Cutler, D., & Song, Z. (2010). Workplace wellness programs can generate savings. *Health Affairs*, 29(2), 304-311.

6. Baker, K. M., Goetzel, R. Z., Pei, X., Weiss, A. J., Bowen, J., Tabrizi, M. J. & Thompson, E. (2008). Using a return-on-investment estimation model to evaluate outcomes from an obesity management worksite health promotion program. *Journal of Occupational and Environmental Medicine*, 50(9), 981-990.

7. Barbosa, C., Bray, J., Brockwood, K., & Reeves, D. (2014). Costs of a work-family intervention: Evidence from the Work, Family, and Health Network *American Journal of Health Promotion, 28*(4), 209-217.

8. Barbosa, C., Bray, J.W., Dowd, W.N., Mills, M.J., Kelly, E.L., Moen, P., & Wipfli, B. (2015). Return on investment of a Work-Family Intervention: Evidence from the Work, Family, and Health Network. *Journal of Occupational and Environmental Medicine*, 57(9), 943–951.

9. Basu, A., & Coe, N. (2015). 2SLS vs 2SRI: Appropriate methods for rare outcomes and/or rare exposures. *Unpublished mansucript, University of Washington, Seattle*.

10. Belotti F., Deb, P., Manning, W.G., Norton, E.C. (2015). Twopm: Two-part models. *The Stata Journal*, 15(1):3-20.

11. Blundell, R., & Dias, M. C. (2009). Alternative approaches to evaluation in empirical microeconomics. *Journal of Human Resources*, *44*(3), 565-640.

12. Box, G., & Cox, D. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society. Series B (Methodological), 26*(2), 211-252. Retrieved from http://www.jstor.org/stable/2984418

13. Bray, J. W., Kelly, E. L., Hammer, L. B., Almeida, D. M., Dearing, J. W., King, R. B., & Buxton, O. M. (2013). An integrative, multilevel, and transdisciplinary research approach to challenges of work, family, and health. Methods report (RTI Press), 1.

14. Buntin, M. B. and Zaslavsky, A. M. (2004). Too much ado about two-part models and transformation? Comparing methods of modeling medicare expenditures. *Journal of Health Economics*, 23: 525–542.

15. Burton J. (2016). WHO Healthy Workplace Framework and Model: Background and Supporting Literature and Practices 2010. Retrieved from http://www.who.int/occupational_health/healthy_workplace_framework.pdf

16. Cameron, A. C., & Trivedi, P. K. (2005). Microeconometrics: Methods and applications.

17. Cameron, A. C., and P. K. Trivedi. (2010). Microeconometrics Using Stata. Rev. ed. College Station, TX: Stata Press.

18. Center for Disease Control and Prevention. (2015, July 7). Wellness at Work. Retrieved January 28, 2017, from https://www.cdc.gov/features/workingwellness/

19. Chapman, C. G. (2014). Identification of population average treatment effects using nonlinear instrumental variables estimators: Another cautionary note. The University of Iowa.

20. De Greef, M., & Van den Broek, K. (2004). Making the case for workplace health promotion. *Analysis of the Effects of WHP*.

21. Dement, J. M., Epling, C., Joyner, J., & Cavanaugh, K. (2015). Impacts of workplace health promotion and wellness programs on health care utilization and costs: results from an academic workplace. *Journal of Occupational and Environmental Medicine*, 57(11), 1159-1169.

22. European Network Workplace Health Promotion. (1997). *Luxembourg declaration on workplace health promotion in the European Union*. Bundesanstalt für Arbeitschutz und Arbeitsmedizin

23. Frick, K. D. (2009). Micro-costing quantity data collection methods. *Medical Care*, 47(7 Suppl 1), S76.

24. King, G., & Zeng, L. (2001). Logistic regression in rare events data. Political analysis, 9(2), 137-163.

25. Gebhardt, D. L., & Crump, C. E. (1990). Employee fitness and wellness programs in the workplace. *American Psychologist*, 45(2), 262.

26. Goetzel, R. Z., & Ozminkowski, R. J. (2008). The health and cost benefits of work site health-promotion programs. *Annual Review of Public Health*, 29, 303-323.

27. Goetzel, R. Z., Pei, X., Tabrizi, M. J., Henke, R. M., Kowlessar, N., Nelson, C. F., & Metz, R. D. (2012). Ten modifiable health risk factors are linked to more than one-fifth of employer-employee health care spending. *Health Affairs*, 31(11), 2474-2484.

28. Goetzel, R. Z., Tabrizi, M., Henke, R. M., Benevent, R., Brockbank, C. V. S., Stinson, K. & Newman, L. S. (2014). Estimating the return on investment from a health risk management program offered to small Colorado-based employers. *Journal of Occupational and Environmental Medicine*, 56(5), 554-560.

29. Goetzel, R. Z., Henke, R. M., Tabrizi, M., Pelletier, K. R., Loeppke, R., Ballard, D. W. & Metz, R. D. (2014). Do workplace health promotion (wellness) programs work? *Journal of Occupational and Environmental Medicine*, 56(9), 927-934.

30. Griffin, S. C., Regan, T. L., Harber, P., Lutz, E. A., Hu, C., Peate, W. F., & Burgess, J. L. (2016). Evaluation of a fitness intervention for new firefighters: injury reduction and economic benefits. *Injury Prevention*, 22(3), 181-188.

31. Gujarati, D. N. (2009). Basic econometrics. Tata McGraw-Hill Education.

32. Hamilton, L. C. (1991). ssi2: Bootstrap programming. Stata Technical Bulletin 4: 18–27. Reprinted in Stata Technical Bulletin Reprints, vol. 1, pp. 208–220. College Station, TX: Stata Press

33. Henke, R. M., Goetzel, R. Z., McHugh, J., & Isaac, F. (2011). Recent experience in health promotion at Johnson & Johnson: lower health spending, strong return on investment. *Health Affairs*, 30(3), 490-499.

34. Horwitz, J. R., Kelly, D. R., DiNardo, J.E. (2013). Wellness incentives in the workplace: Cost savings through cost shifting to unhealthy workers. *Health Affairs*, 32(3), 468-476.

35. Institute of Medicine, (U.S.), & National Academies Press. (U.S.). (2003). Weight Management: State of the Science and Opportunities for Military Programs. Washington, D.C.: National Academies Press.

36. Johns, B., Baltussen, R., & Hutubessy, R. (2003). Programme costs in the economic evaluation of health interventions. *Cost Effectiveness and Resource Allocation*, 1(1), 1.

37. Kaiser Family Foundation (2013, April 25). Summary of the Affordable Care Act. Retrieved January 28, 2017, from http://kff.org/health-reform/fact-sheet/summary-of-the-affordable-care-act/

38. Lewis, A., Khanna, V., and Montrose, S. (2014, November 25). Workplace Wellness Produces No Savings. Retrieved January 28, 2017, from http://healthaffairs.org/blog/2014/11/25/workplace-wellness-produces-no-savings

39. Lewis, A. and Khanna, V. S. (2013, January 16) Is It Time to Re-Examine Workplace Wellness 'Get Well Quick' Schemes? Retrieved January 28, 2017, from http://healthaffairs.org/blog/2013/01/16/is-it-time-to-re-examine-workplace-wellness-get-well-quick-schemes/

40. Hall, K. and Diehm, J. (2013, October 3) Why U.S. Health Care Is Obscenely Expensive, In 12 Charts. Retrieved February 27, 2017, from http://www.huffingtonpost.com/2013/10/03/health-care-costs-_n_3998425.html

41. Maniscalco, P., Lane, R., Welke, M., Mitchell, J. H., & Husting, L. (1999). Decreased rate of back injuries through a wellness program for offshore petroleum employees. *Journal of Occupational and Environmental Medicine*, *41*(9), 813-820.

42. Manning WG, Mullahy J. (2001). Estimating log models: to transform or not to transform? *Journal of Health Economics*, 20:461– 494.

43. Mattke, S., Liu, H., Caloyeras, J., Huang, C. Y., Van Busum, K. R., Khodyakov, D., & Shier, V. (2013). Workplace wellness programs study: Final report. *Rand Health Quarterly*, 3(2).

44. McCoy, M. K., Stinson, M. K., Scott, M. K., Tenney, M. L., & Newman, L. S. (2014). Health promotion in small business: a systematic review of factors influencing adoption and effectiveness of worksite wellness programs. *Journal of Occupational and Environmental Medicine/American College of Occupational and Environmental Medicine*, *56*(6), 579.

45. Mentock, R. (2016. November 14) How Do I Know if I am Being Billed the Correct Amount for My Visit? Retrieved February 27, 2017, from http://familycarepa.com/tag/insurance/

46. Meenan, R. T., Vogt, T. M., Williams, A. E., Stevens, V. J., Albright, C. L., & Nigg, C. (2010). Economic evaluation of a worksite obesity prevention and intervention trial among hotel workers in Hawaii. *Journal of Occupational and Environmental*

Medicine/American College of Occupational and Environmental Medicine, 52(Suppl 1), S8.

47. Merrill, R. M., Hyatt, B., Aldana, S. G., & Kinnersley, D. (2011). Lowering employee health care costs through the healthy lifestyle incentive program. *Journal of Public Health Management and Practice*, 17(3), 225-232.

48. Mossink, J. and Nelson, D. I. (2002). Understanding and performing economic assessments at the company level (No. WHO/SDE/OEH/01.9). Geneva: World Health Organization.

49. Musich, S., White, J., Hartley, S. K., Bhattarai, G. R., Hawkins, K., & Ozminkowski, R. J. (2014). A More Generalizable Method to Evaluate Changes in Health Care Costs with Changes in Health Risks among Employers of All Sizes. *Population Health Management*, 17(5), 297-305.

50. Neira, M. (2013, November). Healthy workplaces: a model for action. In *Presentation at the First Global Healthy Workplaces Conference, London*.

51. Newcomer, K. E., Hatry, H. P., & Wholey, J. S. (2015). Handbook of practical program evaluation. John Wiley & Sons.

52. Nyman, J. A., Abraham, J. M., Jeffery, M. M., & Barleen, N. A. (2012). The effectiveness of a health promotion program after 3 years: evidence from the University of Minnesota. *Medical Care*, 50(9), 772-778.

53. OECD (2017), Health spending (indicator). doi: 10.1787/8643de7e-en (Retrieved March 15, 2017)

54. Quintiliani, L., Sattelmair, J., & Sorensen, G. (2007). The workplace as a setting for interventions to improve diet and promote physical activity. *Documento técnico preparado para el evento conjunto OMS/Foro Económico Mundial sobre la prevención de las enfermedades no transmisibles en el lugar de trabajo]. Ginebra, Organización Mundial de la Salud*

55. Palumbo, M. V., Sikorski, E. A., & Liberty, B. C. (2013). Exploring the cost-effectiveness of unit-based health promotion activities for nurses. *Workplace Health & Safety*, *61*(12), 514-520.

56. Patient Protection and Affordable Care Act, 42 U.S.C. § 10408, (2010).

57. Puhani, P. A. (2012). The treatment effect, the cross difference, and the interaction term in nonlinear "difference-in-differences" models. *Economics Letters*, *115*(1), 85-87.

58. Reardon, J. (1998). The history and impact of worksite wellness. *Nursing Economics*, 16(3), 117.

59. Rush, B., & Ogborne, A. (1991). Program logic models: expanding their role and structure for program planning and evaluation. *The Canadian Journal of Program Evaluation*, 6(2), 95.

60. Sattler, K. M., Deane, F. P., Tapsell, L., & Kelly, P. J. (2018). Gender differences in the relationship of weight-based stigmatization with motivation to exercise and physical activity in overweight individuals. *Health Psychology Open*, *5*(1),

61. Sloan, R. P. (1987). Workplace health promotion: A commentary on the evolution of a paradigm. *Health Education Quarterly*, 14(2), 181-194.

62. Spence, G. B. (2015). Workplace wellbeing programs: if you build it they may NOT come... because it's not what they really need!

63. Stoewen, D. L. (2016). Wellness at work: Building healthy workplaces. *The Canadian Veterinary Journal*, *57*(11), 1188–1190.

64. Terza, J. V., Basu, A., & Rathouz, P. J. (2008). Two-stage residual inclusion estimation: addressing endogeneity in health econometric modeling. *Journal of Health Economics*, *27*(3), 531-543.

65. The National Institute for Occupational Health and Safety (NIOSH) (2012). *Research Compendium: The NIOSH Total Worker HealthTM Program: Seminal Research Papers* 2012. Washington, DC: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, DHHS (NIOSH) Publication No. 2012-146:1-214. Retrieved January 28, 2017, from https://www.cdc.gov/niosh/docs/2012-146/pdfs/2012-146.pdf.

66. The National Center for Health Statistics US. (2016). Health, United States, 2015: with special feature on racial and ethnic health disparities. Retrieved August 24, 2018, from www.cdc.gov/nchs/data/hus/hus15.pdf

67. The National Institute of Diabetes, Digestive, and Kidney Diseases (NIDDK). (2018). Recent advances and emerging opportunities. Retrieved August 31, 2018, from https://www.niddk.nih.gov/about-niddk/strategic-plans-reports/niddk-recent-advances-emerging-opportunities-2018.

68. The National Institute of Diabetes and Digestive and Kidney Diseases. (2018). Understanding adult overweight and obesity. Retrieved Aug. 24, 2018 from https://www.niddk.nih.gov/health-information/weight-management/adult-overweight-obesity/factors-affecting-weight-health.

69. The World Health Organization and The International Commission on Occupational Health (2012, December). Executive summary. *Second International Consultation on Healthy Workplaces, Rome*.

70. Trogdon, J., Finkelstein, E. A., Reyes, M., & Dietz, W. H. (2009). A return-on-investment simulation model of workplace obesity interventions. *Journal of Occupational and Environmental Medicine,* 51(7), 751-758.

71. US Department of Health and Human Services. (2008). 2008 Physical activity guidelines for Americans. Retrieved August 24, 2017, from http://www. health. gov/paguidelines/.

72. Warner, K. E. (1987). Selling health promotion to corporate America: Uses and abuses of the economic argument. *Health Education & Behavior,* 14(1), 39-55.

73. Wholey, J. S. (1983). Evaluation and effective public management. Little, Brown.

74. Widmer, R. J., Allison, T. G., Keane, B., Dallas, A., Lerman, L. O., & Lerman, A. (2014). Using an online, personalized program reduces cardiovascular risk factor profiles in a motivated, adherent population of participants. *American Heart Journal*, 167(1), 93-100.

75. Wooldridge, J. M. (2010). Econometric analysis of cross section and panel data. MIT press.

76. Wooldridge, J. M. (2015). Control function methods in applied econometrics. *Journal of Human Resources*, *50*(2), 420-445.

Appendix 4.A. The First Order Conditions

W_i is the (row) vector of covariates $[period_{it}\ just10_{it}, period_{it}, just10_{it}, X_{it}]$ and θ is the corresponding (column) vector of parameters $[\delta \quad \alpha_1 \quad \alpha_2 \quad \beta]'$.

4.A1. Turnover

The first order conditions for predicted probabilities of turnover with respect to parameters are as follows:

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \sum_{i=1}^{N} turnover_i \left(\frac{\partial \ln(p_i)}{\partial \theta} \right) + (1 - turnover)_i \left(\frac{\partial \ln(1-p_i)}{\partial \theta} \right) = 0 \tag{4.A.1}$$

$$\frac{\partial \ln(p_i)}{\partial \theta} = \frac{\partial p_i}{\partial \theta} \Big/ p_i = \frac{\partial \ln\left(\frac{\exp(W_i\theta)}{1+\exp(W_i\theta)} \right)}{\partial \theta} = \left(\frac{\exp(W_i\theta) W_i}{\left(1+\exp(W_i\theta)\right)^2} \right) \tag{4.A.2}$$

$$= \frac{p_i(1-p_i)W_i}{p_i} = (1-p_i)W_i$$

$$\frac{\partial \ln(1-p_i)}{\partial \theta} = \frac{\partial (1-p_i)}{\partial \theta} \Big/ (1-p_i) = \frac{\partial \ln\left(1 - \frac{\exp(W_i\theta)}{1+\exp(W_i\theta)} \right)}{\partial \theta} \tag{4.A.3}$$

$$= -\left(\frac{\exp(W_i\theta) W_i}{\left(1+\exp(W_i\theta)\right)^2} \right) = \frac{-p_i(1-p_i)W_i}{(1-p_i)} = -p_i W_i$$

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \sum_{i=1}^{N} turnover_i ((1-p_i)W_i) + (1 - turnover_i)(-p_i W_i) = 0$$

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \sum_{i=1}^{N} (turnover_i - p_i) W_i = 0 \tag{4.A.4}$$

where p_i is the predicted probabilities and $\Lambda(W_i\theta)$ is the cumulative density function with $0 < \Lambda(W_i\theta) < 1$. The first order condition (4.A.4) shows that the raw residuals, $(turnover_i - p_i)$, are orthogonal to regressors.

4.A2. Injury

The first order conditions for predicted probabilities of injury with respect to parameters are as follows:

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \sum_{i=1}^{N} injury_i \left(\frac{\partial \ln(p_i)}{\partial \theta} \right) + (1 - injury_i) \left(\frac{\partial \ln(1-p_i)}{\partial \theta} \right) = 0 \tag{4.A.5}$$

$$\frac{\partial \ln(p_i)}{\partial \theta} = \frac{\partial p_i}{\partial \theta} \Big/ p_i = \frac{\partial \ln\left(\frac{\exp(W_i\theta)}{1+\exp(W_i\theta)} \right)}{\partial \theta} = \left(\frac{\exp(W_i\theta) W_i}{\left(1+\exp(W_i\theta)\right)^2} \right) \tag{4.A.6}$$

$$= \frac{p_i(1-p_i)W_i}{p_i} = (1-p_i)W_i$$

$$\frac{\partial \ln(1-p_i)}{\partial \theta} = \frac{\partial (1-p_i)}{\partial \theta} \Big/ (1-p_i) = \frac{\partial \ln\left(1 - \frac{\exp(W_i\theta)}{1+\exp(W_i\theta)} \right)}{\partial \theta} \tag{4.A.7}$$

$$= -\left(\frac{\exp(W_i\theta) W_i}{\left(1+\exp(W_i\theta)\right)^2} \right) = \frac{-p_i(1-p_i)W_i}{(1-p_i)} = -p_i W_i$$

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \sum_{i=1}^{N} injury_i((1-p_i)W_i) + (1-injury_i)(-p_i W_i) = 0$$

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \sum_{i=1}^{N} (injury_i - p_i) W_i = 0 \tag{4.A.8}$$

where p_i is the predicted probabilities and $\Lambda(W_i\theta)$ is the cumulative density function with $0 < \Lambda(W_i\theta) < 1$. The first order condition (4.A.8) shows that the raw residuals, $(injury_i - p_i)$, are orthogonal to regressors.

4.A3. Lost Labor Days

The log-likelihood function is for the Poisson model is

$$\ln L(\theta) = \sum_{i=1}^{N} \ln\left[\frac{\exp(\mu_i)\mu_i^{-lostdays_i}}{lostdays_i!}\right] \tag{4.A.9}$$

$$\ln L(\theta) = \sum_{i=1}^{N} \left[\mu_i - lostdays_i \ln(\mu_i) - \ln(lostdays_i!)\right] \tag{4.A.10}$$

$$\mu_i = \exp(W_i\theta) \tag{4.A.11}$$

where μ_i>0 is the expected value or average of $lostdays_i$

$$\ln L(\theta) = \sum_{i=1}^{N} \left[\exp(W_i\theta) - lostdays_i W_i\theta - \ln(lostdays_i!)\right] \tag{4.A.12}$$

The first order conditions for lost labor days with respect to parameters are as follows:

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \sum_{i=1}^{N} (\exp(W_i\theta) - lostdays_i) W_i = 0 \tag{4.A.13}$$

The first order condition (A2.4) shows that the raw residuals, $(\exp(W_i\theta) - lostdays_i)$, are orthogonal to regressors.

4.A4. Organizational Costs

The log-likelihood function for the two-part model is

$$\begin{aligned}\ln L(\theta) = \sum_{i=1}^{N} & I(cost_i = 0)\ln P(cost_i = 0 \mid x_i) \\ & + I(cost_i > 0)\ln P(cost_i > 0 \mid x_i) f(cost_i \mid cost_i > 0, x_i)\end{aligned} \tag{4.A.14}$$

where *I(A)* is indicator function, equal to 1 if *A* is true, and 0 otherwise.

$I(cost_i = 0) = 1 - I(cost_i > 0)$, without loss of generality.

$P(cost_i = 0 \mid x_i) = 1 - P(cost_i > 0 \mid x_i)$ without loss of generality, is the conditional probability for the first part and estimated using logit model. $f(cost_i \mid cost_i > 0, x_i)$ is the probability density function of the outcome, conditional on the outcome being positive in the second part, which has the gamma distribution ($h(y_i)$) with log link function ($g(\mu_i)$).

$\mu_i = E(y_i) > 0$

$g(\mu_i) = \ln(\mu_i) = W_i\theta$

$$h(y) = \frac{1}{y^{\Gamma}(1/\sigma^2)}\left(\frac{y}{\sigma^2\mu}\right)^{1/\sigma^2} \exp\left(-\frac{y}{\sigma^2\mu}\right), \qquad y > 0 \tag{4.A.15}$$

where *y* represents the outcome for simplification, $\Gamma(k) = (k-1)!$ is the gamma function, and $V(Y_i) = \sigma^2\mu^2$ is the variance of the outcome. Let $I(cost_i > 0)$=1 . The log-likelihood function can be written as:

$$\ln L(\theta) = \sum_{i=1}^{N} (1-I)\ln(1-\Lambda(W_i\theta)) + I\ln\{\Lambda(W_i\theta)g(W_i\gamma)\} \tag{4.A.16}$$

The first order conditions for organizational cost with respect to parameters are as follows:

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \sum_{i=1}^{N} (1-I)\frac{\partial \ln(1-\Lambda(W_i\theta))}{\partial \theta} + I\left\{\frac{\partial \ln \Lambda(W_i\theta)}{\partial \theta} + \frac{\partial \ln g(W_i\theta)}{\partial \theta}\right\} = 0 \tag{4.A.17}$$

$$\frac{\partial \ln(1-\Lambda(W_i\theta))}{\partial \theta} = -\frac{\Lambda(W_i\theta)(1-\Lambda(W_i\theta))W_i}{1-\Lambda(W_i\theta)} = -\Lambda(W_i\theta)W_i \tag{4.A.18}$$

$$\frac{\partial \ln \Lambda(W_i\theta)}{\partial \theta} = -\frac{\Lambda(W_i\theta)(1-\Lambda(W_i\theta))W_i}{1-\Lambda(W_i\theta)} = -\Lambda(W_i\theta)W_i \tag{4.A.19}$$

$$\frac{\partial \ln g(W_i\theta)}{\partial \theta} = \frac{\partial g(W_i\theta)}{\partial \theta} \Big/ g(W_i\theta) = \frac{W_i}{W_i\theta} \tag{4.A.20}$$

where equations 4.A.18 and 4.A.19 are identical procedures to 4.A.2 and 4.A.3. Then,

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \sum_{i=1}^{N} (1-I)(-\Lambda(W_i\theta)W_i) + I\left\{(1-\Lambda(W_i\theta)W_i) + \frac{W_i}{W_i\theta}\right\} \tag{4.A.21}$$

$$= \sum_{i=1}^{N} \left\{ I(cost_i > 0)\left(1 + \frac{1}{W_i\theta}\right) - \Lambda(W_i\theta)\right\} W_i = 0$$

CHAPTER V

CONCLUSION

The idea of promoting health to prevent chronic illnesses has been prevalent for a long time. Take Fielding 1984 and Edington 2014 as examples. These papers are thirty years apart, yet each paper proposes the same strategy: implementing WWPs as solutions to promoting health by targeting the health behaviors of their employees. The consequences of chronic illnesses impact all the stakeholders with treatment cost, income, and productivity losses. The question is whether WWPs achieve their public and private objectives. Despite increasing governmental support for WWP provisions, the effectiveness of WWPs in promoting health continues to be questioned.

The main objective of this dissertation is to offer new approaches for the successful implementation and assessment of WWPs in order to provide rigorous evaluations on the effectiveness of WWPs. The research questions, methods, and findings of this study have significant contributions to the field by answering the following questions:

1. Is research asking the right question when incorporating the value propositions for stakeholders?
2. Can a common impact of WWPs be identified to justify the use of WWPs as robust policy tools in public health?

Our findings show that the narrow focus of economic evaluation literature excludes value propositions of employees and the government. A lack of reliable findings for the program impacts is the result of only focusing on the employers' perspective for outcome evaluations and ignoring the program objectives other than ROI. In addition, there is a potential lack of association between objectives for WWP implementation and program choice where health insurance companies and program vendors might determine the program trends. Expected program outcomes may not be attained due to missing connections between organizational objectives and program choice, as well as the missing alignments of value propositions.

Additionally, our findings present that WWPs differ based on the needs of workplaces, objectives of employers, and other workplace characteristics, such as industry and size. Thus, WWPs are unique to their workplaces and each evaluation can be considered as a single observation. The literature allows us to derive a conclusion mostly for whether large companies have positive ROI. Moreover, we cannot validate whether the evaluated WWPs significantly affect the outcome due to missing statistical inference information.

Current literature does not provide enough reliable evaluations on the effectiveness of WWPs. The field needs rigorous evaluations so that we can discuss the effectiveness of WWPs at the public policy level. The field needs further research that offers evidence on aligning private market incentives for WWP implementation that would lead to efficient use of resources for both public and private perspectives.

Many evaluation studies informed us about possible limitations such as selection bias and data issues. The field is aware of its blind spots, yet fails to deliver solutions to those problems. The first essay, the systematic review, contributes to the field by providing methods that can advance study rigor. Self-selection into program participation that causes positive bias for the program effects is one of the major critiques of WWP evaluations. Even if randomization is the golden standard as the means to remove selection bias, conducting randomized studies is not always feasible for WWPs, especially if new regulations require these programs to be available to all employees and participation is often volunteered. The systematic review proves that the field can improve study quality even if more randomized studies cannot be conducted. The review shows the need for more independent study, more studies in small companies, and improvements in reporting quality and analysis methods.

The articles included in the systematic review provides the objectives of the study but do not discuss the objectives of the organization for implementing WWPs. The second essay, the WELCOA study, verifies that ROI is not necessarily the main objective. ROI could be beneficial to look at the use of an organization's financial resources. Yet it cannot determine the success of the program if ROI is not the objective. In addition, the WELCOA study demonstrated that instead of organizational objectives, marketing trends in workplace wellness might have more influence over the decision-making process. This behavior could prevent organizations from choosing the optimal programs. The systematic review and WELCOA studies both confirm the disconnection

between businesses and academia as well as the disconnection between organizational objective and program choice.

The third essay, the Well-Spring study, evaluates a WWP in a small non-profit company to respond to some of the issues raised in the first and second essays. The Well-Spring study demonstrates how to improve the methodological quality of an observational study with better reporting and analysis methods. The findings show that a small company can produce positive ROI close to the average of large companies that was found in the systematic review. Also, the point estimate of ROI is positive, but not significant, which we cannot derive from the literature. The focus of the third essay is on ROI because of organizational preferences. The company wanted to see whether the program saved any money through reducing the costs of injuries and turnover. However, it is important to note that the Just10 program might not have been the optimal program for Well-Spring to achieve its goals. This essay cannot offer a solution for aligning organizational objectives and program choice.

Not all WWPs are well-designed and well-organized, which affects the evaluations. Thus, identifying best practice designs to assess organizational performance in implementing WWPs seems the right direction for progress. Practitioners and academicians came up with the idea of best practice designs for WWPs. Common benchmark domains are identified as key factors for success. Although economic evaluation literature does not reflect the diversity of program setting and implementation, benchmarks could provide an agreement on the key factors for successful WWP implementations. Best practice designs could provide a framework for concerns related to

the decision-making process and effectiveness of WWPs. The main reasons why evaluations cannot address diversity in program settings and implementation are the missing value propositions of stakeholders and similar relevant data. For example, WELCOA's Well Workplace Checklist does a good job of bringing the major benchmarks together. The checklist provides vague information on the employers' perspective. This information needs to be clear. In addition, information on employee-facing measures and targeted outcomes should be included so we can connect both the organizational goal(s) and performance in implementing WWPs to the program success.

This study contributes to the field by identifying missing alignments in value propositions. This is the first study which uses economic theory to model firm behavior when implementing WWPs. This novel approach is expected to shift why and how we should evaluate WWPs by incorporating organizational values. This study also contributes to the field by providing a rigorous evaluation of a small nonprofit company and sets an example of how to align company objectives for WWP implementation with evaluated outcomes to improve the reporting quality for reliable evidence.