Time Allocation and the Weather

by

Jingye Shi

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Economics

Waterloo, Ontario, Canada, 2012

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The overriding theme of my dissertation is the use of short-term weather fluctuations to study how people allocate their time across activities. In Chapter 1, a theoretical model is developed to distinguish malfeasant from legitimate forms of employee sickness absenteeism. In this model, individuals' marginal utility of indoor leisure is increasing in their sickness levels, while their marginal utility of outdoor leisure is an increasing function of the interaction of their health and the quality of outdoor weather. In equilibrium, sickness absenteeism occurs at both ends of the sickness distribution – among the relatively sick and among the most healthy facing the best weather. The positive relation between marginal changes in weather quality and levels of sickness absenteeism in the workplace reflects the substitution of the inframarginal employees who are the least sick away from work activities towards outdoor leisure activities. The model in Chapter 1 suggests an empirical strategy to identify a shirking component in overall reported sickness absenteeism. Not only does this approach avoid attributing entirely legitimate forms of absenteeism to shirking, but unlike previous studies using employee dismissal rates, it is able to distinguish shirking activity whether or not that activity is detected by employers.

In order to exploit exogenous weather fluctuations to identify shirking activity, we need a one-dimensional measure of weather "quality". The primary objective of Chapter 2 is to construct a weather quality index that captures the influence of the weather on workers' preferences for outdoor leisure activity. The weather quality index takes into account the multifaceted nature of weather conditions, and measures how various weather elements – temperature, humidity, precipitation, wind speed, and cloud cover – come together to affect the propensity of employees to engage in high-utility outdoor recreational activities. The resulting index provides a ranking of different weather conditions in terms of their outdoor recreational values, which can then be used to capture the incentives of employees to shirk contractual work hours in response to purely exogenous weather changes.

Chapter 3 empirically tests the existence of weather-induced substitution between work and outdoor leisure activities and examines how this type of behaviour varies across workers facing different shirking incentives. Linking 12 years of employee data from Canada's monthly Labour Force Survey (LFS), which queries reasons for employees' absences, to weather quality measured using the index constructed in Chapter 2, a clear positive relationship is found between the quality of outside weather conditions and short-term reported sickness absenteeism. Moreover, consistent with a key proposition of the theoretical model in Chapter 1, the empirical relation between weather and sickness absenteeism tends to be larger when existing shirking incentives are low, such as when sick pay is less generous and when probability of getting fired if caught shirking is high. There is, however, little evidence that firms are able to adjust shirking incentives through the payment of efficiency wages.

Finally, Chapter 4 examines another type of substitution induced by weather shocks – the substitution between outdoor and indoor physical activities. The Chapter begins with a theoretical model of the decision to participate in physical activities, which assumes that when adverse weather shocks deter outdoor physical activities, indoor physical activities are the only viable option for individuals to stay physically active. However, because the indoor options are more costly, substituting from outdoor to indoor physical activities is easier for higher-income individuals. This suggests an explanation for the stylized fact that rates of physical activity participation are low among lower socioeconomic groups. Linking time-use data from Canadian General Social Survey with archival weather data, the results of the empirical analysis in this chapter provides evidence of a positive income effect enabling substitution from outdoor to indoor physical activities when outside weather is not conducive for participating in outdoor activities. By exploiting the role that income plays in maintaining physical activity levels when less costly outdoor options are limited, this chapter formally illustrates a credible causal link between people's income levels and their participation in leisure time physical activities and provides direct evidence of this link. The results have important policy implications for promoting physical activities, especially among lower income population.

Acknowledgements

When I started to write these acknowledgements, I realized there are so many people to whom I would like to express my gratitude for their help, in one way or another, during my journey of completing my Ph.D. study.

My sincere gratitude goes first and foremost to my supervisor, Dr. Mikal Skuterud, for his invaluable guidance and support throughout the course of my Ph.D. study. The first three chapters in my dissertation are the products of the research project that Dr. Skuterud and I have been working on together for the past four years. Dr. Skuterud developed the research question, and guided us through the process of model development and empirical analysis, while I was primarily responsible for data preparation and analysis. In the process of my writing up of this dissertation, if it were not for Dr. Skuterud's generous and insightful comments, part of this dissertation would have been quite unreadable. Working under Dr. Skuterud's supervision, I have learned a lot from his immense knowledge, and more important, his enthusiasm for this profession and earnest work attitude for his job have all been such an inspiration to me, which I truly believe will benefit my future career tremendously. As a matter of fact, besides all the things I have learned from him, Dr. Skuterud's incredible patience in putting up with my countless unannounced drop-ins to his office, alone, have made me feel very fortunate to have him as my supervisor.

I would also like to thank all other members of my dissertation committee and external examiner, Dr. Stéphanie Lluis, Dr. Dinghai Xu, and Dr. Daniel Scott for their insightful comments and suggestions, which make this dissertation more accurate and distinctive. Especially, I should thank Dr. Stéphanie Lluis, whose contribution to the project in the beginning has not been forgotten.

Special thanks also go to Brian Mills for his assistance to access the weather data; to our graduate program coordinator, Pat Shaw, for the long six-year consistently considerate assistance she has offered me; and to all the faculty members at the department who have imparted their knowledge to me.

Of course, despite all the wonderful things, doing a Ph.D. definitely involves feeling frustrated and depressed when things were not making sense to me, and that is when I resorted to my friends for a nice home-cooked meal or just talk. Among many of them, I would like to thank two of my dearest friends, Han Wang and Quan Zhou, for all their support and laughter that got me through the rough patches.

Finally, I'd like to express my deepest appreciation to my mother, my deceased father, and my sister, for their unconditional love and endless encouragement in every aspect of my life.

Dedication

To my mother and my deceased father, Jianjun Zhang and Shuxing Shi, for their unconditional love, which has made me want to make them proud.

Table of Contents

List of Tables

List of Figures

Introduction

Time is arguably the most universal, equitable and critical resource constraint that people face. Decisions about how to best allocate one's time are not only paramount in determining people's well-being, but also have important social and economic implications for our society as a whole. For example, the growth of the economy is affected by the time people spend on productive activities; improvements in public health are affected by people's decisions of whether or not to spend time on physical activities; and child development is affected by the time parents devote to their children. Becker's (1965) time allocation theory lays the ground work for using economic theory to study people's time allocation decisions. In his seminal model, people allocate their time to the following activities: sleep, work, leisure, and home production. In a utility-maximizing framework, the marginal utility of the last unit of time spent on each activity must be the same for all activities. The implication of this mechanism underlying optimal time allocation is that the change in the marginal utility of any of these activities will cause utility-maximizing individuals to reallocate their time. Therefore, substitutions away from low-utility activities and towards high-utility activities will occur. Changes in time allocation across activities, in fact, are associated with many social phenomena. For example, overtime work and work absence are associated with substitutions of time use between work and non-work activities; more prevalence of sedentary lifestyles indicate substitution away from active activities and towards more sedentary activities. These social phenomena have important implications for social welfare. Therefore, understanding people's time allocation decisions is important because not only the allocation of time can lead to different social and economic outcomes but also the changes of time allocation offers a unique perspective for understanding people's behavioural decisions behind many social phenomena.

One of the greatest challenges in doing empirical research on time allocation is ruling out the possibility that observed relations are driven by unobservable characteristics that are not controlled for in the analysis. The natural fluctuations in weather conditions, especially short-term weather fluctuations, provide a valuable source of exogenous variation for studying time allocation decisions in the sense that we can be sure that observed relations between weather and time allocation are not being driven by unobservable characteristics. Yet, as an important feature of the environment in which people's daily activities take place, weather is highly relevant in people's decision making process of their time use, because utility associated with daily activities, especially outdoor recreational activities, varies with weather conditions. Weather-induced increases in the marginal utility associated with outdoor recreational activities lead to increase in the demand for these activities, which, given the time constraint, implies there must be less time spent on other activities.

The main theme of my dissertation is the use of exogenous weather fluctuations to study people's time allocation behaviour. Two types of substitutions are of particular interest. First is substitution away from work to leisure when the improvement of outside weather condition increases the utility associated with outdoor leisure activities, which is interpreted as a form of employee shirking behaviour and reported in the data as a short-term absence from work due to one's own illness or the illness of a family member. Exploiting this particular type of substitution in time allocation provides an approach for understanding the phenomenon of workplace absenteeism. Specifically, it informs the nature of workplace absenteeism by clearly distinguishing malfeasant absence from legitimate forms of absence, which is critical for designing effective interventions to reduce workplace absenteeism. The other type of substitution examined is substitution away from outdoor to indoor physical activities when outdoor weather is not conducive for outdoor physical activities. Understanding the factors that influence this type of substitution, such as the income levels of individuals, has important policy implications for promoting healthy lifestyles in some particular segment of the population, which can in turn reduce the health disparities in society.

Chapter 1 proposes a theory of shirking sickness absenteeism by modelling the substitution of employees from work activities to outdoor leisure activities in response to fluctuations in outdoor weather conditions. In this model, the marginal utility of leisure of employees depends not only on their level of sickness, but also on the weather. Since the type of high-utility outdoor recreational activities that employees are likely to substitute towards when the weather improves tend to be more enjoyable when one is healthy, employees' marginal utility of outdoor leisure is expected to be decreasing in sickness. In equilibrium, sickness absenteeism in this model occurs at both ends of the sickness distribution – among the relatively sick and among the most healthy facing the best weather, but the positive relation between weather quality and sickness absenteeism only reflects the behaviour of the inframarginal employees who are the least sick, which the model interprets as shirking. Moreover, the model shows that employees' response to shirking incentives changes significantly with the wage offered to them, which indicates employers can potentially manage employees' malfeasant absence behaviour through optimal wage setting. The findings of this model not only provide insights for understanding absenteeism behaviour, particularly malfeasant absenteeism, but also provide an empirical strategy for directly capturing shirking behaviour, which is reflected in the correlation between weather and sickness absenteeism.

In order to take advantage of exogenous variation in the weather conditions in studying individuals' time allocation decisions, weather quality needs to be measured quantitatively. The second chapter of my dissertation is devoted to constructing a weather quality index, which provides a scalar measure of the value individuals attach to weather conditions. An average person's weather preference is revealed by how much he or she is willing to be exposed to a particular type of weather. To do this, time-use data from the Canadian General Social Survey (GSS) is linked to archival weather data, allowing us to empirically relate the actual activities people engage in to the multifaceted weather conditions they face at precisely the same point in time. Then, average weather preferences in the population are captured by identifying how various weather elements – temperature, humidity, precipitation, wind speed, and cloud cover – come together to affect the propensity of an employee with a regular daytime schedule engaging in high-utility outdoor recreational activities on their weekends or weekday evenings. The resulting index, therefore, takes values between 0 and 1 representing the preferences for recreation-related weather conditions. Hence, the weather index constructed in Chapter 2 not only provides a ranking of different weather conditions in terms of their values for doing outdoor recreation, but also directly measures the incentives to substitute away from other activities, such as contractual work hour obligations.

Chapter 3 empirically tests the existence of weather-induced changes in time allocation between work and leisure activities, which captures a type of shirking behaviour, and examines how the change varies across workers facing different shirking incentives. Linking 12 years of employee data from Canada's monthly Labour Force Survey (LFS), which queries reasons for employees' absences, to weather quality measured by the index constructed in Chapter 2, the correlation between employees' short-term sickness absence and weather quality is examined. A positive and statistically significant relationship in the non-winter months is found between the quality of outside weather condition and sickness absenteeism.[1] Based on the logic of the theoretical model in Chapter 1, it is argued that this correlation between weather and sickness absence captures a shirking component out of the overall reported sickness absenteeism. Moreover, comparing this relation across workers facing different shirking incentives, such as between salaried employees, who are typically paid for time off, and hourly paid employees who are not, we found employees facing high existing shirking incentives are less responsive to weather improvements. However, little evidence that firms are able to adjust shirking incentives through the payment of efficiency wages are found.

In Chapter 4, I turn to examine another particular type of substitution induced by weather fluctuations, which is the substitution between outdoor and indoor physical activities. First, a theoretical model is developed to show that when adverse weather conditions deter outdoor physical activities, indoor physical activities are the only viable option for

[1]Non-winter months include April, May, June, July, August, September, and October.

individuals to participate in physical activity. However, because the indoor options are more costly, substituting from outdoor to indoor physical activities is easier when people have less financial constraints. Therefore, if one of the underlying causes of higher income leading to more physical activities is providing a larger set of activity options, participation in physical activities is less sensitive to weather for people with higher income, and the difference between physical activity participation across income levels is larger, when weather condition is not conducive for outdoor activities. Linking GSS time-use data with archival weather data, results of the empirical analysis provide evidence of the hypothesized positive income effect enabling substitution from outdoor to indoor physical activities when outside weather deters people from participating in outdoor activities. By exploiting the role that exogenous weather fluctuations play in influencing physical activity decisions, this chapter generates a credible causal link between people's income levels and their participation in leisure time physical activities, which has important policy implications for promoting physical activities, especially among lower socioeconomic groups.

Relative utilities of the daily activities to which people allocate their time to can be significantly altered by outside weather conditions, mainly through the outdoor activities they engage in. Reallocating time across these activities not only has direct social and economic impacts on various aspects of our society, but also reflects many social phenomena through the underlying substitution process behind the time reallocation. Using weather variations to understand the causal factors of these phenomena is important for improving the wellbeing of our society.

Chapter 1

An Efficiency Wage Model of Shirking Sickness Absenteeism

1.1 Introduction

According to the Canadian Labour Force Survey (LFS), a total of 32.9 billion hours were scheduled to be worked in Canada over the 2011 calendar year. Of these hours, 11.9% were lost to employee absences. In comparison, over the same period of time, lost hours due to unemployment, based on the reported hours preferences of job seekers, amounted to 7.3% of total scheduled hours.[1] In contrast to the social and academic attention the unemployment issue has drawn, employee absence has been relatively understudied, despite its potential cost to the economy. For example, the American Economic Association's electronic bibliography EconLit contains 17,473 published articles containing the word "unemployment" in its title or abstract, while the word "absenteeism", in comparison, appears in only 284. Although the loss in economic activity caused by employee absences can be made up through replacement workers or overtime, to the extent that replacement workers cannot be found, overtime can not fully cover the lost work hours, or there are spillovers in lost production (e.g., in an assembly operation or were work is organized in teams), the cost of employee absences can be substantial. Some employee absences, such as vacation and statutory holidays, are scheduled, and their effects on the organizations are relatively easier to absorb. Other absences, for instance those caused by illness, can

[1]Based on my own calculation using the 2011 LFS public-use microdata files. Total scheduled hours are total usual weekly hours (all jobs) of all employed workers across the 12 monthly LFS files of 2011. To get an annual value, this total is multiplied by 52/12. Absent hours are either full week or part week for any reasons other than holidays and vacations. Preferred weekly hours of the unemployed are based on their preferences for full-time or part-time work and mean usual weekly hours of currently employed full-time and part-time workers.

not be planned ahead, so such absences are more disruptive to the production process and more costly to the economy as a whole. For policy makers or employers finding the most effective policies to reduce employee absences, it is of great importance to understand the underlying causes of this workplace phenomenon.

The vast majority of unscheduled absences, both short- and long-term, are reported by statistical agencies and employers to be health related. Polls, however, suggest that deliberately misreporting actual health to avoid contractual work hours is commonplace. For example, a recent poll by Careerbuilder.com found that one-third of U.S. workers called in to work sick with a fake excuse at least once in the past year (*Globe and Mail*, October 24, 2008). Which policies are most effective in reducing sickness absenteeism depends critically on the extent to which absenteeism is unavoidable, and therefore legitimate. Studies have suggested that employers' investment in promoting employees' health can be considered as an investment in human capital, and their profits can be improved from the decrease in the health care services utilized by an employee or his/her family as well as from the increase in worker productivity (Goetzel and Ozminkowski, 2008; Nicholson et al., 2005). Therefore, to the extent that sickness absenteeism reflects genuine sickness, policies would seem to be best served promoting worker health. Employers, in this case, might even want to encourage the take up of sick days. A study by Chatterji and Tilley (2002) actually shows that sick employees' work attendance, sometimes called "presenteeism", have negative impacts on productivity, so employers should consider increasing sick pay in order to discourage sick employees showing up at work. Evidence that a substantial proportion of observed sickness absenteeism is malfeasant in nature would, alternatively, point to policies that enable employers to better monitor true worker health or increase the cost to workers of using sick days to reduce work hours.

Regardless of differences in the underlying motivations of employees, legitimate and malfeasant sickness absenteeism are difficult to be distinguished from each other. Information regarding employees' true health is highly asymmetric between employers and employees. This asymmetry is exactly what provides employees with an incentive to shirk. Moreover, even in the situation when true health information is available, such as to employees themselves and their doctors, distinguishing the two types of sickness absenteeism is still difficult since sickness is not measured on a single dimension that can be easily compared to some threshold level that employers or policy makers have predetermined and made explicit in company policy.

Acknowledging this inherent difficulty of defining and identifying malfeasant sickness absenteeism, this chapter proposes an efficiency wage model to capture a shirking component in sickness absenteeism and examines the effect of firms' wage setting behaviour on this employee behaviour. In the model, employees' marginal utility of leisure depends not only on their level of sickness, but also on outdoor weather conditions. Since the type of high-utility outdoor recreational activities, such as golfing and bicycling, that employees are likely to substitute towards when the weather improves tend to be more enjoyable when one is healthy, employees' marginal utility of outdoor leisure is assumed to be decreasing in sickness. In equilibrium, sickness absenteeism in this model occurs at both ends of the sickness distribution – among the relatively sick and among the most healthy facing the best weather. More important, comparative statics analysis shows that only the sickness absenteeism of the most healthy workers increases with marginal improvements in weather quality. Given the inherent ambiguity of distinguishing legitimate from illegitimate sickness levels for taking sick absence, we argue that the relation between weather and sickness absenteeism, which only occurs among the most healthy workers, more clearly reflects behaviour that is malfeasant in nature. The model predicts that when firms do not act to affect employees' shirking behaviour, the relation between weather and sickness absenteeism is larger when existing shirking incentives are low, for example when workers are not paid for time off work. However, when firms exert some control over workers' behaviour by offering efficiency wages, workers' optimal responses to existing shirking incentives change substantially. This model not only precisely identifies a shirking component of the overall sickness absenteeism and highlights an efficiency wage mechanism through which wage rates may be used as a method for controlling work absences, but also provides a strategy for identifying shirking activities in empirical studies.

The following section of the chapter will provide a more complete review of the existing literature that studies absenteeism behaviour across disciplines. Section 3 then presents a version of the efficiency wage model to study absenteeism with a focus on shirking absenteeism. Section 4 discusses the key implications of the model, followed by section 5 that summarizes the findings of the proposed model, and concludes the chapter.

1.2 Related Literature

Absenteeism has been most extensively studied in the fields of organizational/industrial psychology and management science. The psychology literatures are focused primarily on the individual approach, which emphasizes that employees' work absenteeism is determined internally within the individual employee and influenced by both the individual's psychological, such as personality, and demographic characteristics. In one of the seminal contributions to absenteeism theory, psychologists Steers and Rhodes (1978) propose the *process model* of employee absence, which is based on the individual approach and describes how individual characteristics, such as personality, education, health status, financial status and demographic characteristics indirectly affect absenteeism, through a set of medial variables including individuals' job attitude, job satisfaction, and ability to attend work. Building on this framework, Rosse and Miller's (1984) *theory of job adaptation* and Hulin, Roznowski and Hachiya's (1985) *withdraw theory* explain absenteeism as individuals' behavioural responses to changes in the medial variables mentioned above. Absenteeism, especially malfeasant absenteeism, is a voluntary decision an individual makes. In this sense, the individual approach psychologists usually take offers valuable insights for understanding the individual's decision making process regarding absenteeism. However, the focus of these studies is mainly about employees being physically absent from work, regardless of the reasons. Therefore, the theories they proposed to explain people's decisions regarding work absence ignore the differences in people's motivations between malfeasant and legitimate absence.

In comparison to the psychology literatures, the management science literatures concerned with absenteeism are more focused on the social context of absenteeism, rather than considering it solely as private behaviour. In these literatures, organizational characteristics, such as the absenteeism culture at the workplace (Johns and Nicholson, 1982), employers' monitoring effort or permissiveness level for absenteeism (Brooke, 1986), and flexible work hour arrangement (Allen and Shockley, 2009), are argued to have influential impacts on employees' absenteeism. Although the effects of some of the organizational characteristics, such as employer monitoring effort, are implicitly assumed to be imposed on absence that is malfeasant in nature, these literatures still lack a unifying theory that clearly distinguishes malfeasant from legitimate forms of absenteeism.

The earliest theoretical models of absenteeism in the economics literature similarly overlooked the distinction between legitimate and malfeasant absenteeism (Allen 1983; Dunn and Youngblood 1986; Barmby, Orme and Treble 1995). In theses models, absenteeism is viewed as an optimal labour supply response by utility-maximizing employees to contractual hours constraints imposed by employers. Assuming certain limitations on the ability of workers to costlessly change jobs, anything in this framework that serves to increase a workers's marginal utility of leisure, be it the flu, an unpleasant work environment, or a sunny day, will create incentives for workers to absent themselves from work. But considering work absence solely as the consequences of employees' labour supploy decisions overlooks both optimal firm behaviour in balancing malfeasant absenteeism and presenteeism, by for example providing workers with a finite number of paid sick days each year, and the risk inherent in an employee's decision to misreport health. As a result, these literatures fail to explain why the incidence of malfeasant absenteeism varies across labour markets. For example, why do rates of absenteeism tend to be higher in the public sector and in unionized workplaces (Cunninghan, Debben, and James, 2001; Nielsen, 2008; Vandenheuvel, 1994; Wooden, 1990).

Theories that do not consider the effects of firm behaviour on absenteeism provide little or no insight into how employers can most effectively adjust employee incentives to minimize malfeasant absence. Asymmetric information refers to a situation in which one party in a bilateral relationship has more or superior information than the other. This information asymmetry can cause the party that lacks information to not be able to detect a breach of the contract, which in turn gives one party an incentive to not fulfill the contract. In the context of labour market, if workers' work efforts are difficult to measure or observe, information regarding workers' work effort would be a classic example of asymmetric information. In this case, asymmetric information between employers and employees regarding the effort levels of employees gives employees incentives to shirk. In this situation, profit-seeking firms might be willing to pay a higher wage to increase the cost of job loss and in turn to reduce the shirking incentives, which means the higher wage can be used to prevent workers' shirking behaviour.

Economists Carlin (1989) and Barmby, Sessions and Tremble (1994) borrow this idea from the canonical efficiency wage model of Shapiro and Stiglitz (1984) to define a unifying theoretical relationship between legitimate and malfeasant sickness absenteeism. They argue that in workplaces where workers' true health status is imperfectly monitored, the threat of firing if caught misreporting health status can be an effective device to enforce worker honesty and reduce malfeasant sickness absenteeism, if the cost of job loss is sufficient. In both papers, firms raise wage rates above the market-clearing wage, which has two effects. First, it increases the cost of job loss, since now the employee faces the possibility of losing a job paying a higher wage than what is being offered in the rest of the market. Second, to the extent that this wage-setting behaviour is optimal, all employers in the market raise wages above the market-clearing, which results in equilibrium unemployment, so that employees who are dismissed for shirking are no longer assured of finding a replacement job in the following period. As a result, employees' expected cost of illegitimately using sick days to reduce work hours is higher in terms of both a higher wage loss if dismissed and a reduced probability of finding a replacement job. Therefore, employers can use increasing wage rates to reduce employees' misuse of sick days, which is a type of malfeasant absenteeism. In contrast to the no-shirking equilibrium of Shapiro and Stiglitz, wage rates in these two studies are set so as to accept some optimal level of shirking, which depends on exogenous factors in the model, such as the amount of sick pay offered or the cost of absenteeism to employers, in terms of lost production. The existence of some level of shirking at the equilibrium offers an explanation of the real-world phenomenon of malfeasant sickness absenteeism.

Studies using efficiency wage theory to explain variations in sickness absenteeism across organizations implicitly recognize the difference between malfeasant and legitimate sickness absenteeism, because absence from work due to genuine health issues are arguably less sensitive to wages. However, these studies remain somewhat arbitrary in the sense that the approach used for determining the legitimacy of an incidence of sickness absenteeism is still comparing the employee's sickness level to a threshold level of sickness that often is arbitrarily set by policy makers or employers. Following the notion that the malfeasant sickness absenteeism is likely more responsive to exogenous factors other than true health, the model in this chapter attempts to use employees' responses to the plausible exogenous weather variations to identify a component in the overall sickness absenteeism that is unambiguously malfeasant in nature.

We can think of at least three reasons why malfeasant absenteeism will respond to the weather. First, there now exists considerable evidence linking stock market values to the weather, which is attributed to the effect of sunshine on investor moods and, in turn, preferences for risk (e.g., Saunders 1993; Kamstra, Kramer and Levi 2003; Kliger and Levy 2003; Levy and Galili 2008; Jaconsen and Marquering 2008). Since shirking is similarly a decision involving risk, it is reasonable to expect employees to be more willing to shirk when weather conditions are good. Second, exposure to good weather may directly affect people's moods in a beneficial way, as evidence from the psychology literature suggests (e.g., Howarth and Hoffman 1984). However, the mechanism we have in mind in this chapter is through a third channel. In particular, we assume that when weather conditions are good, workers' marginal utility of outdoor leisure will tend to increase, because good weather enables particular outdoor activities that either give people pleasure or makes these activities more enjoyable. Assuming that these weather improvements are valued most by the most healthy employees, since they are best able to exploit the conditions in their recreational activities, we are able to clearly distinguish a component in overall absenteeism that is malfeasant in nature. Using this theoretical construct, we can then begin to examine how various company policies affecting shirking incentives influence levels of reported absenteeism in equilibrium.

1.3 Model

The model of shirking sickness absenteeism that follows builds on the model of Barmby, Sessions and Treble (1994). In their static model, an individual's utility is an increasing function of leisure and income, with the individual attaching a weight to each depending on his or her own health level. Barmby, Session and Treble (1994)'s model is extended in this study in two directions. First, employees' marginal utility of leisure depends not only on their level of sickness, but also on outdoor weather conditions. Second, rather than modelling work absence as a one-time decision, we assume that the employees are infinitely-lived agents, so that the risk in shirking also entails the possibility of beginning the following period without a job to go to. Therefore, shirking decisions are also a function of job acquisitions rates, which we believe resembles the real-world risk in misreporting health to an employer more closely.

In any period, we assume ex-ante identical risk-neutral individuals receive utility $U = (1-\delta)y + \delta(T-h)$, where T is a time endowment; h are hours worked; and y is income. In making labor supply decisions, individuals weigh the relative marginal utility of leisure spent outdoors and indoors. When spent outdoors, we assume $\delta = (1-\theta)\,\lambda$, where θ reflects an individual's level of sickness and λ is an index of weather quality. In contrast, when spent indoors, the marginal utility of leisure is assumed to be independent of the weather, but is increasing in sickness, specifically $\delta = \theta$. An individual, therefore, prefers outdoor to indoor leisure if $\theta < \lambda/(1+\lambda)$, where the two state-dependent parameters, θ and λ, are assumed randomly (uniform) and independently distributed in the population over the interval [0,1].

Individuals receiving an employment contract, who opt to satisfy the contractual hours obligation h, receive wage w. Employees who choose not to show up for work, on the other hand, and whose true sickness level is either legitimate or goes undetected, receive sick pay $s < w$. The threshold sickness level beyond which absence is deemed legitimate is given exogenously to employees by θ^z. However, θ^z can be thought of as being determined endogenously by the employer as it trades off the costs of absenteeism among healthy and productive employees and what Chatterji and Tilley (2002) refer to as the "presenteeism" of unhealthy, unproductive, and perhaps also contagious employees. The employer's technology for monitoring employee sickness detects an individual's true sickness level θ with probability α at cost k, which is sufficiently small, so the technology is always employed. In the event that illegitimate absence ($\theta < \theta^z$) is detected, a shirking employee is not only dismissed and forced to sustain himself on an unemployment benefit $b < s$ in the current period, but must also begin the following period unemployed facing an exogenous job acquisition rate $a < 1$.

Given this setting, the lifetime utility of an infinitely-lived individual beginning period one with an employment contract can be written:

$$U = \begin{cases} U^{na} = (1-\delta)\,w + \delta\,(T-h) + \rho\,V(E), \text{if not absent in period 1} \\ U^{a} = (1-\delta)\,s + \delta\,T + \rho\,V(E), \text{if absent and not dismissed in period 1} \\ U^{u} = (1-\delta)\,b + \delta\,T + \rho\,a\,V(E) + \rho\,(1-a)\,V(U), \text{if absent and dismissed in period 1} \end{cases}$$

where $\rho\,\epsilon[0,1]$ is a time preference discount rate and $V(E)$ and $V(U)$ are continuation values from period 2 forwards if beginning period 2 with or without a contract, respectively.

Defining $E(U^e)$ and $E(U^u)$ as the expected utilities (over the distributions of θ and λ) of being employed and unemployed in any period, respectively, $E(U^e)$ is always larger than $E(U^u)$. This is a because when an individual begins a period being unemployed, she will receive the utility of $U = (1-\delta)b + \delta\,T$ with certainty. However, when the individual is employed at the beginning of a period, the expected utility she will receive in that period is a probabilistic mixture of being unemployed (in the case of getting caught shirking), employed but absent from work, and employed at work. Individuals who start the period employed and whose true sickness level is above the threshold θ^z can take legitimate sickness absence and receive utility $U = (1-\delta)s + \delta\,T$. Since sick pay s is higher than the unemployment benefit b, the utility these individuals are guaranteed to receive is higher than the unemployed. For people who start the period employed and whose true sickness level is below the threshold θ^z, the worst that can happen is being caught shirking when they illegitimately take sick leave, so the utility they will receive during the period will be at least as good as starting the period unemployed. Given $[E(U^e) - E(U^u)] > 0$, the continuation value from period 2 forwards, if beginning period 2 with a contract, is greater than beginning period 2 without a contract. The difference is: [2]

$$\begin{aligned} V(E) - V(U) &= \sum_{t=1}^{\infty} \rho \left[1 - \alpha\,(\theta^o + \theta^z - \theta^i) - a\right]^{t-1} [E(U^e) - E(U^u)] \\ &= \frac{1}{1 - \rho\,[1 - \alpha\,(\theta^o + \theta^z - \theta^i) - a]}\,[E(U^e) - E(U^u)] > 0. \end{aligned}$$

Therefore, in deciding whether to shirk the contractual work obligation h in the first period, employees not only take into account the risk of a lower income level b in the current period, but also that they are always better off beginning the next period in the employed state, whether or not they choose to be absent in that period. [3]

The expected lifetime utility of an illegitimately ill employee who chooses to shirk is $U^s = \alpha\,U^u + (1-\alpha)\,U^a$. Shirking occurs if $U^s > U^{na}$, which defines a threshold for the marginal utility of leisure given by:

$$\delta^c = \frac{w - \alpha\,b - (1-\alpha)\,s + \rho\,(1-a)\,[V(E) - V(U)]}{w - \alpha\,b - (1-\alpha)\,s + h} \tag{1.3.1}$$

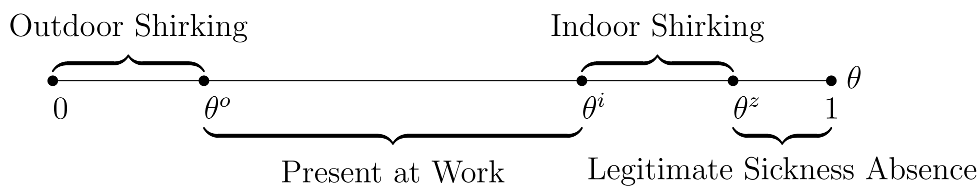
beyond which employees prefer to be absent from work. A worker who prefers outdoor leisure will, therefore, choose to be absent if sickness lies below the outdoor sickness threshold $\theta^o = (\lambda - \delta^c)/\lambda$, while a worker preferring indoor leisure will choose absence if sickness exceeds the indoor sickness threshold $\theta^i = \delta^c$. As long as $\theta^i > \theta^o$, so that at least someone shows up for work, we are insured that $\lambda - \delta^c - \lambda\,\delta^c < 0$ or $\theta^i > \lambda/(1+\lambda) > \theta^o$. The proportion of employees who shirk is then $(\theta^o) + (\theta^z - \theta^i)$, where the first and second term

[2] The formal proof is presented in Appendix A.

[3] The model implies that individuals always accept the contracts offered by employers, which may not be the case if employers use other contract structures to reduce absenteeism. For example, if employers request an individual to post a bond when accepting the job offer and the bond could be withheld in the case of the individual being caught having a malfeasant absence (see Carmichael (1985) for a discussion of such a contract structure), the person's job acceptance decision depends on the sequence of the events assumed in the model. If the individual has information on θ and λ, given the amount of bond requested, the person makes the decision by comparing the utility of not accepting the offer with the higher value of U^{na} and U^s. However, if the realization of the two state-dependent parameters, θ and λ, occurs after the individual's job acceptance decision, the job acceptance decision is made by comparing the utility of not accepting the offer with the higher value of expected U^{na} and U^s. In both situations, once the job offer is taken, an individual's decision-making process regarding absence remains the same as in the model with no employee bond, but the difference between $V(E)$ and $V(U)$ is larger, which results in a higher threshold value of δ^c.

capture outside and inside shirking absenteeism, respectively. This is presented in Figure 1.

Figure 1: Worker's optimal labour market choice over the distribution of sickness level



Notes: Sickness level, θ is uniformly distributed between 0 and 1 with 0 indicating perfect health.

Since the weather, unlike all the remaining exogenous variables of the model, only affects the outdoor sickness threshold θ^o, one can readily see that any relation between the weather and reported sickness absenteeism, must reflect the behavior of the inframarginal employees who are the most healthy. Given the practical difficulty of determining the legitimacy of sickness in the vicinity of θ^z, the observed positive relation between the weather and sickness absenteeism here more clearly reflects malfeasant absence behaviour than where sickness falls marginally below the arbitrary threshold θ^z. This leads to the first proposition.

Proposition 1 (Weather-absenteeism relation) *Marginal improvements in outdoor weather quality above some critical level $\lambda = \delta^c$ lead to an increase in sickness absenteeism that is unambiguously illegitimate.*

Proof. The expected proportion of employees $\bar{n}$ who accept an employment contract and choose to shirk is $\Pr(\theta < \theta^o) + \Pr(\theta^i < \theta < \theta^z) = (\theta^o) + (\theta^z - \theta^i)$, given that θ is uniformly distributed over the positive unit interval. But since θ^i depends only on the threshold marginal utility of leisure δ^c, given by equation (1.3.1), and θ^z is an exogenous constant, a marginal improvement in the weather λ only affects the extent of outdoor shirking θ^o. Given that $\theta^o = (\lambda - \delta^c)/\lambda$, we have:

$$\frac{\partial \Pr(\theta < \theta^o)}{\partial \lambda} = \begin{cases} 0, & \text{if } \lambda \leq \delta^c \\ \delta^c/\lambda^2, & \text{if } \lambda > \delta^c \end{cases} \tag{1.3.2}$$

which implies reported sickness absenteeism is a discontinuous increasing function of the weather.

Although the weather-absenteeism relation is necessarily positive, its magnitude does vary with other shirking incentives. Specifically, if existing shirking incentives are low, such as where the risk of being detected shirking is high or the job acquisition rate is low, inframarginal non-shirking workers who are induced to shirk by the weather improvement are relatively healthy, compared to the situation where the shirking incentives are already high. But since the influence of weather on the marginal utility of outdoor leisure is largest for employees who are the most healthy, it is in these workplaces where existing outdoor shirking levels are lowest, that the weather has its biggest impact. This leads to the second proposition.

Proposition 2 (Interaction effects in partial-equilibrium model) *The marginal effect of the weather on sickness absenteeism is larger where existing shirking incentives are low, that is where the threshold marginal utility of leisure for choosing to be absent from work (δ^c) is high. This implies that we should see a larger weather-absenteeism relation: (i) where job acquisition rates are low; (ii) where sick pay is less generous; and (iii) where the probability of being dismissed when shirking is high.*

Proof. The marginal effect of the weather on sickness absenteeism is given by $\partial\theta^o/\partial\lambda = \delta^c/\lambda^2$. Given the exogeneity of weather λ, applying the implicit function theorem to equation (1.3.1), we have:

$$\frac{\partial(\frac{\partial\theta^o}{\partial\lambda})}{\partial a} = \frac{1}{\lambda^2}\frac{\partial\delta^c}{\partial a} = \frac{\left(\frac{dw}{da}\right)[h-\rho\alpha(1-a)V]-\rho\alpha V(d+h)}{(d+h)^2} \tag{1.3.3}$$

$$\frac{\partial(\frac{\partial\theta^o}{\partial\lambda})}{\partial s} = \frac{1}{\lambda^2}\frac{\partial\delta^c}{\partial s} = \frac{\left(\frac{dw}{ds}-(1-\alpha)\right)(h-\rho\alpha(1-a)V)}{(d+h)^2} \tag{1.3.4}$$

$$\frac{\partial(\frac{\partial\theta^o}{\partial\lambda})}{\partial \alpha} = \frac{1}{\lambda^2}\frac{\partial\delta^c}{\partial \alpha} = \frac{\left(\frac{dw}{d\alpha}+s-b\right)[h-\rho\alpha(1-a)V]+\rho(1-a)V(d+h)}{(d+h)^2} \tag{1.3.5}$$

where $V = V(E) - V(U)$ and $d = w - \alpha b - (1-\alpha)s$. The assumed distributions of the sickness and weather parameters insure that $\delta^c < 1$, which implies that $\rho\,(1-a)\,[V(E)-V(U)] = \rho\,(1-a)\,V < h$. Since we know $V > 0$ and $h - \rho\alpha(1-a)V > 0$, and that $d > 0$ (since $w > s > b$), in the absence of any employer wage adjustments (wage derivatives are zero), the signs of all three derivatives are unambiguous: $\partial(\partial\theta^o/\partial\lambda)/\partial a < 0$; $\partial(\partial\theta^o/\partial\lambda)/\partial s < 0$; and $\partial(\partial\theta^o/\partial\lambda)/\partial\alpha > 0$.

Because $\delta^c < 1$, it can be shown that $\partial\delta^c/\partial w > 0$. The wage rate, therefore, provides employers with an instrument to control the extent of shirking absenteeism. To model employers' optimal wage-setting behavior, we assume the following sequence of events: (i) employers randomly offer $\bar{n}$ contracts providing wage w for h hours of work; (ii) all received offers are accepted (since $s > b$ and $V(E)-V(U) > 0$); and (iii) individuals receive sickness

(θ) and weather (λ) realizations and decide whether or not to attend work. This implies that although employers have just as much information about the current period's weather as their employees, they are unable to use the weather to influence wage-setting, since they must set wages prior to the weather being realized. As a result, Proposition 1 holds whether one thinks about the partial- or full-equilibrium model, that is, whether efficiency wages are paid or not. However, as we show below, the marginal effect of other shirking incentive parameters on the weather-absenteeism relation can be sharply different if the firm is able to use the wage to influence shirking incentives.

Turning to the full-equilibrium model, suppose firm revenue is an increasing and concave function of the number of contracted employees who turn up for work, given by $R(n)$ (normalizing the price of output to one). Assuming employers know the unconditional distribution of the weather, and that there is no outdoor absenteeism when $\lambda \leq \delta^c$, their expected profits in any period are:

$$\begin{aligned} \mathrm{E}\left\{\pi(w,\bar{n})\right\} &= \int_0^{\delta^c} R(\theta^i \bar{n})\, d\lambda + \int_{\delta^c}^1 R\left((\theta^i - \theta^o)\bar{n}\right)\, d\lambda - \\ & \int_0^{\delta^c} \left[\left(\theta^i w + (1-\theta^i)k + (1-\theta^z)s + (1-\alpha)(\theta^z - \theta^i)s\right)\bar{n}\right] d\lambda - \\ & \int_{\delta^c}^1 [(\theta^i - \theta^o)w + (1-\theta^i+\theta^o))k + (1-\theta^z)s + (1-\alpha)(\theta^z - \theta^i + \theta^o)s)\bar{n}]\, d\lambda \end{aligned}$$

where the four cost terms within both pairs of square parentheses reflect the costs associated with wages, monitoring, legitimate sick pay, and illegitimate sick pay, respectively. Solving the integrals, replacing the sickness thresholds with δ^c and simplifying, the employer's problem amounts to choosing $\bar{n}$ and w to maximize:

$$\mathrm{E}(\pi) = R(\Delta^c\,\bar{n}) - (\Delta^c(w - k - (1-\alpha)s) + (1-\alpha\theta^z)s + k)\,\bar{n} \qquad (1.3.6)$$

subject to $\Delta^c = 2\delta^c - 1 - \delta^c \log(\delta^c)$ and equation (1.3.1). Since w enters the profit function both directly and through Δ^c (via δ^c), the firm faces the usual efficiency-wage model tradeoff in setting the wage, between lowering labour costs directly and limiting shirking incentives. The solution yields the two first-order conditions:

$$\frac{\partial \mathrm{E}(\pi)}{\partial w} = R'(\cdot)\,\frac{\partial \Delta^c}{\partial w}\,\bar{n} - \left[(w-k-(1-\alpha)s)\frac{\partial \Delta^c}{\partial w} + \Delta^c\right]\bar{n} = 0 \qquad (1.3.7)$$

$$\frac{\partial \mathrm{E}(\pi)}{\partial \bar{n}} = R'(\cdot)\,\Delta^c - [\Delta^c\ (w-k-(1-\alpha)s) + (1-\alpha\,\theta^z)s + k] = 0 \qquad (1.3.8)$$

which together implicitly define the equilibrium wage rate:

$$(\Delta^c)^2 - \frac{\partial \Delta^c}{\partial w}\,[(1-\alpha\,\theta^z)s + k] = 0. \qquad (1.3.9)$$

Total differentiation of equation (1.3.9), to identify employer wage adjustments, together with equation (1.3.1), reveals that employer wage responses vary dramatically across other shirking incentive parameters. In particular, optimal wage adjustments are larger when both shirking incentives and direct costs increase, as in the case of an increase in sick pay s, than when only shirking incentives are affected, as in the case of the job acquisition rate a.[4] In fact, assuming efficiency wages are paid, an exogenous increase in sick pay s, leads employers to more than fully adjust wages, so that shirking incentives in the new equilibrium are actually lower. This distinction yields the final proposition.

Proposition 3 (Interaction effects in full-equilibrium model) *If employers augment the shirking incentives of employees through wage adjustments, that is pay efficiency wages, the weather-absenteeism relation is: (i) unambiguously decreasing in the job acquisition rate, but is attenuated relative to the partial-equilibrium case; (ii) unambiguously increasing in the generosity of sick pay; and (iii) either increasing or decreasing in the probability of being dismissed when shirking.*

Proof. Applying the implicit function theorem to the solution of the profit maximization problem in equation (1.3.9) and using equation (1.3.1) to identify optimal employee responses to these adjustments yields:

$$\frac{\partial w}{\partial a} = \frac{\rho\,\alpha\,V(d+h)\,[f+2\Delta^c\,(d+h)]}{2\,[h-(1-a)\rho\,\alpha\,V]\,[f+\Delta^c\,(d+h)]} \tag{1.3.10}$$

$$\frac{\partial w}{\partial s} = \frac{2(1-\alpha)f+2\Delta^c(1-\alpha)(d+h)+(1-\alpha\theta^z)(d+h)}{2\,[f+\Delta^c(d+h)]} \tag{1.3.11}$$

$$\frac{\partial w}{\partial \alpha} = \frac{-\,[2(s-b)(f+\Delta^c d)+d\theta^z s]}{2\,[f+\Delta^c(d+h)]} \tag{1.3.12}$$

where $f=(1-\alpha\theta^z)s+k$. Substituting these optimal wage responses into equations (1.3.3), (1.3.4),and (1.3.5), yields:

$$\frac{\partial(\frac{\partial\theta^o}{\partial\lambda})}{\partial a} = \frac{1}{\lambda^2}\frac{\partial\delta^c}{\partial a} = \frac{-\,\rho\,\alpha\,V f}{2\,[f+\Delta^c(d+h)]\,(d+h)} < 0 \tag{1.3.13}$$

$$\frac{\partial(\frac{\partial\theta^o}{\partial\lambda})}{\partial s} = \frac{1}{\lambda^2}\frac{\partial\delta^c}{\partial s} = \frac{(1-\alpha\theta^z)\,[h-\rho\,\alpha\,(1-a)V]}{2\,[f+\Delta^c(d+h)]\,(d+h)} > 0 \tag{1.3.14}$$

$$\frac{\partial(\frac{\partial\theta^o}{\partial\lambda})}{\partial \alpha} = \frac{1}{\lambda^2}\frac{\partial\delta^c}{\partial \alpha} = \frac{-\,\theta^z s\,h+(2-\theta^z s\,\alpha)\rho(1-a)V}{2\,[f+\Delta^c(d+h)]\,(d+h)} \gtreqless 0. \tag{1.3.15}$$

[4]To see this, note that both s and a enter equation (1.3.1), but only s enters the firm's profit function, given by equation (1.3.6).

Comparing the predictions in the partial- and full-equilibrium models, malfeasant sickness absenteeism, which is reflected in the correlation between sickness absence and the weather, responds to shirking incentives differently when profit-maximizing employers use wages to manage employees' shirking incentives. The theoretical predication that employees' shirking decisions are affected by firm wage setting behaviour is consistent with efficiency wage theory, which implies that employers can use wage-setting to control employee sickness absenteeism.

1.4 Discussion

The model's prediction that employees who are already facing higher shirking incentives are less responsive to weather improvements is somewhat counterintuitive. In general, we would expect to see people who are less likely to be fired for shirking or more likely to receive a very generous sick pay to be more likely to use sickness as an excuse for their absence from work to exploit nice weather. However, this is not the case in our model. The model provides a theoretical explanation for why the popular perception about malfeasant absenteeism, or employee shirking behaviour in general, may be misleading. When employees face low shirking incentives, the inframarginal workers who will be induced to take malfeasant sickness absence for enjoying nice outdoor weather are relatively healthy, compared with the situation when employees already face high incentives and many healthy workers have already been using sickness absence illegitimately. Because healthier individuals enjoy outdoor leisure activities more, these employees – the healthier inframarginal employees who face lower existing shirking incentives – will be more responsive to weather improvements. This counterintuitive implication of the model suggests a strategy for using the real world data to test the general incentive structure of our model, which can inform us the reliability of the model for explaining illegitimate absenteeism as well as other employee shirking behaviour in a real world setting.

In existing studies, the distinction between legitimate and illegitimate absence is vague in both the conceptual and empirical work studying absenteeism behaviour. This is not only because the information regarding employees' genuine health is asymmetric between employees and employers, but the arbitrarily set sickness threshold for determining the legitimacy of sickness absence – θ^z in our model – causes issues. For example, in a situation in which two individuals both have a sickness level in the neighbourhood of θ^z, where one is slightly below the θ^z, while the other one is just above. In this case, the needs of the two individuals for taking a day off to convalesce may be very similar. However, according to the rule, their use of absence could be viewed as behaviour that is fundamentally different, where one is legitimate and the other one is malfeasant. Therefore, relying on a subjective standard to determine the legitimacy of sickness absence seems problematic. In comparison, the employee behaviour underlying the absence-weather correlations in this model is arguably much more clearly malfeasant in nature, regardless of the threshold level of θ^z. The model not only theoretically captures a particular type of malfeasant absenteeism, but provides a strategy to empirically identify this shirking component out of the overall sickness absenteeism.

In the model, employees who take sickness absence are at both ends of the sickness distribution, so that these employees are both the least and most healthy in the workforce. Among the relatively sick employees taking sickness absence, some will be deemed as taking an illegitimate day off work, since their sickness level will not be high enough relative to the legitimate threshold. The behaviour of these employees is treated in the same way as those who are perfectly healthy but still taking sickness absence. Therefore, even among the employees who are identified as illegitimately taking sickness absenteeism, the nature or underlying motivation of their behaviours are fundamentally different. At the upper end of the sickness distribution, the reason for absence is ultimately related to health issues. On the other hand, employees from the lower end of the distribution are missing work for completely different reasons unrelated to health, in the case considered here, to exploit nice weather. However, one could think of any number of reasons driving up the marginal utility of leisure away from work, such as an important sporting event, having a similar effect.

For employees who are relatively sick but below the threshold for taking a legitimate sickness absence, one could also argue that their productivity is likely to suffer from their low health level. Consequently, reducing the absence of these people to increase work attendance may actually have negative impacts on workplace productivity. In a recent global survey on workplace absence in 8 countries – Australia, Canada, China, France, India, Mexico, the U.K. and the U.S. – conducted by the corporate consulting company Kronos Incorporated, the overwhelming reason employees reported for calling in sick to work when they are not actually sick, was high stress levels. Considering the impact of mental health on individuals' performance at work, allowing these employees to take days off might be beneficial for employers. Therefore, employers designing policies to reduce malfeasant absenteeism should recognize the differences in the nature of employees taking absence, and inappropriate policies can actually do harm to their organizations. For example, more severe penalties could reduce the incidence of healthy employees taking absence to enjoy favourable weather conditions, but it can also push sick and unproductive employees to show up at work. An alternative strategy that would reduce opportunistic absenteeism, but have no impact on health-related, whether mental health or otherwise, absenteeism, is to offer employees a flexible work schedule that allows them to coordinate their work schedules more freely. This could reduce employee incentives to use work hours for other purposes, without inducing employees who are truly sick to show up at work. Therefore, effective policies aimed at reducing malfeasant absenteeism should really target those who are perfectly healthy and only use sickness absenteeism as an excuse for shirking, and be cautious about encouraging employees with health issues to attend work.

1.5 Summary

This chapter proposes a model, in which legitimate and malfeasant sickness absenteeism are clearly distinguished without depending on an arbitrarily determined standard for legitimate sick leave. The model relies on two critical assumptions. First, improvements in weather quality can significantly increase the utilities associated with outdoor leisure activities. Second, healthier people derive more utility from engaging in outdoor leisure activities. With these two assumptions, the model demonstrates that nice weather actually induces workers who are relatively healthy to take sickness absences to enjoy the outdoor leisure activities. This behaviour is clearly a form of shirking, which ought to be treated differently from legitimate sickness absence when policies aim to reduce workplace absenteeism.

The model also demonstrates that employers can effectively influence employees' absenteeism behaviour and control absence by augmenting shirking incentives through wage-setting. When employers passively deal with sickness absence, the nice weather induces more malfeasant sickness absenteeism when existing shirking incentives are low. In contrast, if profit-maximizing employers actively set wage rates to change employees' cost of getting caught shirking, the effects of shirking incentives on the weather-absenteeism correlations are attenuated and in the case of sick pay reversed. This model is able to precisely capture the shirking component in sickness absenteeism, and demonstrates that employees' malfeasant sickness absenteeism can indeed be affected by the wage offered to them, which complies with the efficiency wage theory. The finding in this model not only provides insights for understanding employees' absenteeism behaviour, particularly for malfeasant absenteeism, but also provides a strategy for directly identifying shirking activities in empirical studies.

Chapter 2

A Weather Quality Index for Outdoor Recreational Activities

2.1 Introduction

Among the countless resource constraints limiting economic decisions, time is unique in its equitable allocation – everybody has 24 hours in a day and 7 days in a week. From long-term decisions, such as the optimal number of years invested in education, to short-term decisions, such as hours or minutes spent watching television, decisions about how best to allocate time are paramount in determining people's well-being. From a macroeconomic point of view, individual time-allocation decisions also have profound impacts on society as a whole. For example, the growth of the economy is affected by the time people spent on work; improvements in public health are affected by people's decisions of whether or not to spend time on physical activities; and child development is affected by the time parents devote to their children.

The study of time allocation is concerned with how individuals allocate their scarce time to different tasks and activities. Using economic theory to understand time allocation decisions, individuals are assumed to spend time in those activities that, on the margin, provide the largest utility gains, give financial constraints. Therefore, any factors that affect the utility associated with these activities can be expected to influence time allocation decisions. The weather as an important feature of the environment, in which many of people's daily activities take place, is one such factor affecting time allocation decisions.

Without question, some activities can only be engaged in, or are more enjoyable, under particular weather conditions, which implies that utility gains associated with some activities are more sensitive to weather conditions than others. For example, spending time at the beach or playing tennis outside is more enjoyable on warm sunny days than on cold rainy days. In other words, the demand for certain activities can change dramatically with weather conditions. Geographers have long been interested in the effect of climate change on demand for tourism, since many tourists are motivated primarily by climatic considerations and even those whose motives for travel are definitely non-climatic, such as in education or cultural tourism, would still have interests in selecting times of the year when weather conditions are most suitable. When demand for some activities increases, time allocated to other activities (at least one) must necessarily decrease, due to the time constraint. Chapters 3 and 4 examine the substitutions between activities. Chapter 3 is focused on substitution between time spent on work-related activities and leisure activities, whereas Chapter 4 is focused on substitution between outdoor and indoor leisure activities. The similarity of Chapter 3 and 4 is that they both examine substitutions induced by weather conditions.

The appeal of using weather conditions in studying people's time allocation decisions is its exogeneity in the sense that people cannot influence outdoor weather condition they face. A ubiquitous complication in studying the determinants of human decisions is ruling out the possibility that observed relations are, in fact, driven by unobservable characteristics that are not controlled for in the analysis. Weather shocks, however, especially short-term weather fluctuations provide researchers with precious exogenous variations in the sense that we can be sure observed relations between the weather and time allocation decisions are not being driven by unobservable characteristics that are correlated with the weather. However, in order to take advantage of the exogenous nature of weather conditions, we need a way to measure the weather.

At one level, empirically identifying weather preferences in a population is no different than identifying preferences for ordinary consumption goods. When someone choose to expose themselves to the weather by going outdoors, they are making a choice to consume the weather. As with nearly all consumption goods, preferences for weather in the population are individual specific. Some people, for example, like exceptionally hot weather more than others, just like some people, but not others, enjoy Brussels sprouts. In identifying preferences empirically, the best that can be done is to identify average preferences in the population. Of course, part of what makes preferences individual-specific is that people value activities differently. Someone who likes to read books will tend to have much weaker preferences for weather than someone who prefers to spend her leisure time gardening or sailing. But, of course, this is also the case for many final consumption goods, in which the ultimate use of the good varies across individuals. For example, someone who uses a bicycle to get to work will tend to have stronger preferences for bicycles than someone who only uses them for recreation. Also similar to most consumption goods, housing being a particularly good comparison, the weather is multifaceted with its various dimensions, such as rain and cloud cover, tending to be highly collinear. In estimating weather preferences in a population, one therefore has to identify the marginal effect of a given weather element, holding the other dimensions of the weather constant, as is done in any hedonic pricing model.

What makes the weather very different from other consumption goods, however, is that there is no direct market for the weather, in which equilibrium prices and quantities bought and sold can be used to infer consumer preferences. The reason is that the weather is a non-excludable public good provided freely by nature. The closest one can get to a market for weather would be to exploit price variations for real estate or vacations across geographic locations offering different expected weather. But, of course, isolating the effect of weather from other differences across locations is far from straightforward. An alternative approach is to identify preferences off behaviour in a situation where there is clearly no market, that is where weather is free.

The objective of this chapter is to construct a weather quality index representing average weather preferences for outdoor recreational activities. The construction of this index is primarily motivated by the objective of empirically examining the theoretical propositions of Chapter 1 (which is the focus of Chapter 3). As the shirking model of sickness absenteeism in Chapter 1 assumes, weather quality can significantly influence the utility employees obtain from outdoor leisure activities, leading employees who work indoors to substitute away from contractual work hours to exploit desirable weather. To capture this idea, the weather index constructed in this chapter focuses on a broad set of high-utility outdoor leisure activities, and estimates how an individual's utility of engaging in these outdoor recreational activities is affected by various weather elements, which is empirically captured in the propensity of engaging in these outdoor leisure activities.

To construct the weather quality index, Canadian General Social Survey (GSS) time-use data, which identifies the detailed activities of survey respondents continuously over a randomly assigned 24-hour period, is merged with Environment Canada's hourly weather data on five weather elements – temperature, relative humidity, precipitation, wind speed, and cloud cover – in 56 Canadian cities in non-winter months. By merging these two datasets, the activities of respondents at the top of each hour are linked with local weather conditions these respondents face at precisely the same point in time. Since we want to see how weather influences people's decisions to be outdoors or indoors, as much as possible we want to identify this behaviour at a time when people's time allocation choices are not constrained. We, therefore, restrict attention to the weekend and weekday evening activities of paid employees with regular daytime work schedules.

The results from the analysis show that warmer weather results in a higher propensity for outdoor leisure activities up to some threshold temperature, the value of which depends on the amount of wind. The "bliss point" combination of weather condition is a humidex of 27.2 ℃, a wind speed of 14.7km/hr and clear skies. Physical weather conditions (rain and/or high wind speed) negate the effects of humidex and wind speed, resulting in less participation in outdoor leisure activities, hence, lower weather qualities. Two types of robustness analysis have also been conducted by excluding activities that could potentially bias the estimates of the index, such as the leisure activities with ambiguous locations and home production activities, and the results do not change substantially.

In the following section of the chapter, we review the existing literature that attempts to estimate weather preferences, as well as the economics literature that exploits the exogeneity of the weather for identification. Section 3 introduces the activity and weather data employed in the empirical analysis, and section 4 discusses the empirical specification used and related analysis conducted in this study. In section 5, the results of the constructed weather index are presented and discussed, and an example of using the index is demonstrated. Section 6 summarizes the findings and discusses the potential uses of this weather quality index.

2.2 Related Literature

The effect of weather and climate conditions on agriculture is probably the topic that has received the most attention in economics literature. Moschini and Hennessy (2001) once said in the *Handbook of Agricultural Economics*: "uncontrollable elements, such as weather, play a fundamental role in agricultural production."(p.89) Researchers have used variations in agricultural productions caused by exogenous weather variations to study the production and consumption decisions of economic agents. For example, Paxson (1992) uses regional rainfall in Thailand to construct estimates of shocks to transitory income of Thailand farm households, which are then used to study these household's saving behaviour. Roll (1984) uses the effects of temperature and rainfall on the output of orange, and in turn on the price change of orange juice, to study the demand function of concentrated orange juice. Similarly, Angrist, Graddy and Imbens (2000) use wind-generated waves at sea as an instrumental variable to identify the price elasticity of the demand for fish.

Outside of this agriculture literature, much of the effects of weather on human behaviour have been interpreted as evidence of the psychological effects of the weather. Saunders (1993) attributes the small but significantly positive correlation between sunshine and the stock prices to investors' good mood on sunny days, a phenomenon of which there is evidence in the psychology literature. Starr-McCluer (2000) examines the effect of weather on retail sales and finds a modest but significant role for unusual weather in explaining monthly fluctuations in sales. One of the explanations she gives for her findings is that the effect of unusual weather on consumers' moods can temporarily affect their demand for goods, but this effect of weather tends to be washed out at a quarterly frequency.

Compared to the literature using weather to understand production and consumption decisions, studies explicitly focused on the effect of weather on people's time-use decisions are much less common. Most of the existing studies in this realm are either about the effect of weather on time spent on work-related activities or about the effect of weather on the decisions regarding choices of leisure activities, especially choices made between sedentary and active leisure time activities. Connolly (2008) relates daily hours of work to the incidence of rain and finds that rain yesterday reduces time at work today by 25 and 33 minutes for men and women, respectively, which provides evidence of intertemporal labour supply substitution responses to weather.

Linking U.S. time-use data to daily weather data between 2003 and 2006, Zivin and Neidell (2010) find a significant reduction in labour supply in climate-exposed industries as temperature increase beyond 30 ℃. They also examine the effect of weather on leisure activities and find that time spent on outdoor leisure activities at temperatures below -5 ℃ is 37 minutes less than when temperature is between 24 ℃ and 26 ℃, while indoor leisure decreases by roughly 30 minutes between these two temperature ranges. Their findings again indicate the potential substitution of time use across activities. The significant findings in Zivin and Neidell's study regarding the response in leisure activities to weather conditions are not surprising. In fact, a large body of studies in health economics have paid much attention to the effect of weather on physical activity behaviour. Strong evidence that higher temperature and less rain induce people to substitute away from sedentary leisure activities, such as watching T.V., towards participating in physical activities exists (Merill et al, 2005). Moreover, there is also evidence of substitution between outdoor and indoor leisure-time physical activities caused by weather changes (Eisenberg and Okeke, 2009).

Although the effects of weather on various aspects of daily life have been examined, the challenge of measuring weather preferences has not received much attention from economists. In the empirical studies examining weather effect or that use the effect of weather to derive variations in other factors, various weather elements are sometimes included directly in the analysis, but more often only a single weather element, such as temperature or rain, is considered. To the extent that time allocation decisions are affected by more than a single element, this approach, by not utilizing all available information about the weather, is inefficient. The primary motivation for constructing a weather quality index is to examine its correlation with sickness absenteeism in order make inferences regarding employee shirking behaviour. Therefore, the weather quality index constructed in this chapter aims not only to represent the people's rankings of weather conditions, but also to directly measure the propensity of people to engage in outdoor leisure activities as a function of these weather conditions.

Mieczkowski (1985) is among the first to attempt to construct a composite measure of climatic well-being of tourists. Based on a general knowledge about the influence of climatic conditions on the physical well-being of humans, including preferred temperatures, the role of relative humidity, and the cooling influence of the wind, Mieczkowski (1985) proposes the "Tourism Climate Index" (TCI). The TCI measures suitability of weather conditions for general tourism activities, considering thermal sensation (which is related to skin temperature and affected by other factors, such as relative humidity and wind speed), wind speed, rainfall and sunshine. Using a scale of 0 to 5, in half-unit increments, to measure the optimality of weather variables, the TCI is a linear and separable function: $TCI = 4cid + cia + 2R + 2S + W$. In the formula, cid represents the daytime comfort index, composed of maximum daily dry bulb temperature and minimum daily relative humidity; cia is daily comfort index, composed of mean daily dry bulb temperature and mean daily relative humidity; and R, S, and W denote precipitation, daily hours of bright sunshine and wind speed, respectively. TCI is a comprehensive index in the sense that it provides a method to systematically evaluate the climate resource associated with tourism for locations around the world. However, there are three main limitations of applying TCI practice. First, the index is designed for an average tourist, so it may not be useful for evaluating climate conditions for specific tourist activities. Second, the assessment of each weather element is based on an adaptation of previous studies on thermal comfort and lacks validation in the field. Third, the weights attached to each weather variable in the index function are highly subjective.

Recognizing these issues, attempts have been made in later studies to improve it. Morgan et al. (2000) rank the relative importance of the weather elements based on in situ surveys when rating climate conditions for 3S (sea, sand and sun) tourism. Scott and McBoyle (2001) modify the TCI function of Mieczkowski (1985) by replacing the temperature term with a more recent measure of temperature called apparent temperature and use the resulting measure to examine the impact of projected climate change on tourism demand in a sample of tourism destinations in North America. Last, in their study developing a second generation climate index for tourism (CIT), De Freitas, Scott and McBoyle (2008) estimate weather preferences for beach activities using a sample of students, where they not only consider the effects of various aspects of weather, but also include an "overriding" effect of physical weather conditions, rain and strong wind. Although all of these studies have in their own ways improved the reliability of TCI, relying on surveys to identify weather preferences implies these weather indices might still be ultimately subjective and arbitrary.

In contrast with studies discussed above, which identify ideal weather conditions for tourism activities using stated preferences, this study uses revealed preferences in time-use data to construct a weather quality index. Not only is there a strong tradition in economics of using revealed choices to make inference about preferences, which dates back to the seminal work of Paul Samuelson (1938), but also we believe revealed preferences are especially advantageous in evaluating weather preferences. In examining the effect of weather on time allocation – particularly time spent at work – the weather quality index we are trying to construct intends to capture a broad set of outdoor activities, not just one or a few particular tourism-related activities. When identifying weather preferences regarding a broad set of activities, survey questions asking respondents about hypothetical weather conditions become more problematic as responses may depend on the activities respondents have in mind. In addition, we would like the weather quality index to have some transparent interpretation rather than simply providing a ranking of weather conditions. We construct our index by modeling how various weather elements come together to jointly influence the probability of people engaging in outdoor recreational activities.

2.3 Data

The basic empirical strategy employed in this chapter is to link the weather condition an individual faces to the activity he or she is engaged in. Specifically, we use data from Environment Canada's National Climate Data and Information Archive (NCDIA), which provides hourly interval weather data on five elements: temperature, relative humidity, precipitation, wind speed, and cloud cover. The hourly weather data are always collected at the top of the hour. We have obtained these data, spanning 30 years or more, from weather stations located in 56 of Canada's Census Metropolitan Areas (CMAs) and Census Agglomerations (CAs). To estimate the weather index, the hourly NCDIA weather data is merged with time-use data from three waves (1992, 1998, and 2005) of Statistics Canada's General Social Survey (GSS). The time-use data identify the detailed activities of survey respondents continuously over a 24-hour period.[1] By linking time-use data with the hourly weather data, we are able to link the activities of respondents at the top of each hour with local weather conditions at precisely the same point in time.

We restrict our sample to respondents surveyed between April and October for two reasons. First, common outdoor recreational activities conducted during non-winter months are very different from winter time. Because weather preferences are not only individual-specific, but also activity-specific, weather preferences during the non-winter months should be modeled separately from the winter. Second, the motivation for constructing this weather index is ultimately to use it to identify shirking absenteeism, reflected in the correlation between weather and reported absenteeism. However, during the winter months, the absenteeism-weather correlation may reflect factors, such as inclement weather conditions, that have nothing to do with shirking incentives. By focusing on non-winter months, or more accurately the non-winter months, we can avoid this potential source of bias. The analysis is also restricted to the weekend and weekday evening observations of people who are employed full-time and have regular day work shifts to insure that their decisions about

[1]Time-use data records the occurrence and duration information regarding 181 detailed activities. These activities are grouped into 9 categories: employed work; domestic work; care giving for household members; shopping and services; personal care; school and education; organizational, voluntary and religious activity; entertainment; sports and hobbies; and media and communication.

where to spend the time are free from labour market considerations. Extracting the weekend (9am-9pm) and weekday evening (5pm-9pm) records between April 1 and October 31, we are left with a final sample of 33,908 observations on 5,686 wage and salary workers currently employed in a job with regular full-time daytime schedule.

One of the limitations of the GSS data is that it does not directly identify whether activities are indoors or outdoors. Using our own judgement, we categorize all recreational activities according to their locations into three groups: "clearly outdoor", "clearly indoor", and "ambiguous". We identify 13 leisure activities as "clearly outdoor", which are: gardening; walking, hiking, jogging, or running; golf; fairs, festivals, circuses or parades; bicycling; pleasure drives; fishing; boating; rowing, canoeing, kayaking, wind surfing or sailing; zoos; camping; hunting; horseback riding, rodeo, jumping, and dressage. Our analysis is primarily focused on these outdoor leisure activities, because these activities are most responsive to outside weather and could potentially, given the right weather conditions, provide participants with sufficient utility gains to substitute away from other activities, such as contractual work hours obligations, which are the focus of the next chapter. Table 2.1 shows the detailed activity breakdowns of the sample. Together these outdoor leisure activities account for 6.1% of the 33,908 top-of-hour weekend and weekday evening observations in our sample.

The time-use survey also queries respondents as to which of all the activities they engaged in over the 24-hour period they "enjoyed most." Using this information, we can estimate the probability of identifying a particular activity as the most enjoyed, conditional on a respondent at some point engaging in that activity, separately for each activity in the data. Taking the average of these probabilities over sets of activities, there is clear evidence that the set of 13 outdoor activities we have identified do indeed tend to capture the types of high-utility activities we are interested in. Specifically as shown in Table 2.1, among the 13 activities we define as outdoors, they are on average the most enjoyed in 41.0% of cases. In comparison, the average probability among recreational activities that are indoors, such as watching television, is only 16.1%.[2]

Weather elements used in constructing the index are: temperature, relative humidity, precipitation, wind speed, and cloud cover. According to the geography literature, these five weather elements are the most relevant factors when people evaluate weather quality (Mieczkowski, 1985; De Freitas, Scott, and McBoyle, 2008). De Freitas, Scott and McBoyle (2008) distinguish between three facets of the weather: thermal, aesthetic and physical. The thermal facet of weather refers to variables that affect people's sensation of the weather being "hot" or "cold", and is composed of temperature and relative humidity in this study.

[2]Note that, in averaging the probabilities over activities, we weight activities by their relative incidence. For example, every incident of horseback riding in the data is identified as the most enjoyed activity in the day, but there are a trivial number of observations on this activity, so that it contributes essentially nothing to the average probability.

In fact, the humidex, which is similar to the heat index widely reported in the U.S., is developed by J.M. Masterton and F.A. Richardson of Canada's Atmospheric Environment Service in 1979 and used to represent the thermal aspect of weather in this study. The formula used to calculate humidex is given by

$$h = T + \frac{5}{9} \cdot \left(\frac{6.112 \cdot 10^{\frac{7.5 \cdot T}{237.7+T}} \cdot H}{100} - 10 \right)$$

which is a combination of temperature, T (℃), and relative humidity, H (%). Since humidex is only suitable for capturing the thermal sensation in the higher temperature range, we only replace temperature with humidex when temperature is higher than 15℃. The aesthetic is captured by the proportion of the sky covered by cloud, which is recorded in the data on a 10-point scale. The physical facet of weather refers to precipitation or wind. De Freitas, Scott and McBoyle (2008) proposed a theory that physical weather elements tend to nullify the effect of other weather elements. For example, if it is raining, the temperature should have little or no influence on the decision of whether or not to go to the beach. Following this idea, in this study the weather is defined to have a physical condition that can override the effect of other elements, when there exists any precipitation and/or a wind speed in excess of 38km/hour. The wind speed threshold is determined based on the Beaufort Scale, which is an empirical measure that relates wind speed to observed conditions at sea or on land. The threshold of 38km/hour corresponds to 8 or a "Strong Breeze" on this Scale. At this speed: "Large tree branches are set in motion; whistling is heard in overhead wires; umbrella use becomes difficult; and empty plastic garbage cans tip over."

Using 33,908 sample observations, Figure 2.1 plots estimated density functions of weather variables.[3] Compared with temperature, the humidex appears to have a fat-tailed distribution, suggesting that relative humidity has a greater influence on people's thermal sensation at higher temperatures. Fewer than 1% of the observations in our data have wind speed in excess of our threshold of 38 km/hr, which is reflected in the thin-tailed wind distribution beyond 38km/hr. Precipitation accounts for most incidence of physical conditions. Together, precipitation and wind speed account for 12% of the total observations with physical conditions.

2.4 Empirical Specification

Although it is well accepted that the effect of weather conditions on time allocation, whether tourism-related activities or a broader set of outdoor recreational activities, comes

[3]The plots are smoothed with epanechnikov kernel weighting functions.

from the joint effects of multiple weather elements, how exactly these elements come together to affect decisions is not clear. Therefore, the main challenge this study faces when estimating the effect of weather elements is to determine the correct functional form of this weather index.

To obtain some preliminary sense of what the main patterns are, the analysis begins with nonparametric regression analysis to explore how individual weather elements are correlated with the probability of engaging in outdoor recreational activities. Following on the idea of an overriding effect of physical weather conditions proposed by De Freitas, Scott and McBoyle (2008), the analysis is conducted separately for the observations with physical weather conditions and the observations without. Our expectation is that the effect of other weather elements, such as humidex, will be attenuated in the sample of observations with rain or strong wind.

Four weather variables – humidex, physical weather condition (rain or strong wind), wind speed, and cloud cover – are included in the specification for estimating the weather index. Because the nonparametric function identifies single-peaked functions in the effect of wind and humidex, the basic function form of the weather index includes a quadratic form of these two variables. Also, based on the theory of an overriding effect of physical weather condition (De Freitas, Scott and McBoyle, 2008), which is also supported by the nonparametric analysis, the effect of other weather variables will be restricted to zero when there exists physical weather conditions. Therefore, weather quality index is estimated by a probit regression:

$$Prob(outdoors_{ict} = 1) = \Phi(v_{ict}) = \int_{-\infty}^{v_{ict}} 2\pi^{-1/2} exp(-v_{ict}^2/2)\, dv_{ict} \tag{2.4.1}$$

where v_{ict} is specified as follows:

$$\begin{aligned} v_{ict} = \beta_0 + \beta_1 p_{ct} + (1 - p_{ct}) \cdot [\alpha_1 h_{ct} + \alpha_2 h_{ct}^2 + \alpha_3 w_{ct} + \alpha_4 w_{ct}^2 + \\ \alpha_5 (h_{ct} * w_{ict}) + \alpha_6 (h_{ct}^2 * w_{ct}) + \alpha_7 d_{ct} + \mathbf{z}_c \gamma + \mathbf{x}_t \delta] \end{aligned} \tag{2.4.2}$$

where $outdoors_{ict}$ is a dummy variable indicating individual i, residing in city c, at hour t is engaged in an outdoor activity; p_{ct} is a dummy indicating physical conditions; h_{ct}, w_{ct}, and d_{ct} are the humidex, wind speed and cloud cover, respectively; $\mathbf{z}_c$ is a row vector of city dummies; and $\mathbf{x}_t$ is a vector of month (April to October) and hour dummies (9am-9pm).

The functional form is specified to be a nonlinear relationship between outdoor activities and humidex because we expect, and see some evidence in the nonparametric analysis, that the relationship is positive in lower humidex ranges and negative in higher ranges. In other words, exceptionally cold and hot thermal sensation are less attractive for people participation in outdoor activities, as compared to moderate values. The interaction terms of humidex and wind are included to capture the cooling effects of the wind. Although weather is an exogenous factor, to the extent people can influence the weather they face by choosing where they reside, the exogeneity of weather is questionable. For example, people who have relatively strong preferences for warm weather, may also be the kinds of people who have strong preferences for physical activity. Comparing weather and activities of people across cities, we will the tend to see warm weather associated with higher levels of physical activity. But, of course, this does not mean that warm weather increases the likelihood of a given person engaging in physical activities, the effect that we are interested in identifying. To deal with this potential source of bias, we include a full set of 56 city fixed effects. The time variables – such as month and hour of the day – are controlled for similar reasons. These variables are not only associated with people's activity patterns, but also correlated with weather. For example, July and August are two months in which people are most likely to engage in the outdoor recreational activities, in part because this is when weather is most conducive to these activities, but also because this is when people are more likely to be on vacations and therefore have more time available to engage in outdoor recreation. For this reason, we would not want to necessarily attribute weather patterns in July and August to weather preferences. Similarly, park opening hours, which may be very seasonal, may create spurious correlations between the seasonal weather and activities, but it again does not identify people's weather preferences. Therefore, it is critical to condition the analysis on where individuals live, as well as the month of the year and the hour of the day. Once these factors are controlled for, the weather is plausibly orthogonal to any unobservable individual heterogeneity, so that we can interpret the marginal weather effects as pure causal effects, even in the absence of any demographic control variables.[4]

Although we control for the effects of month of the year and hour of the day, we do not condition on day of the week. Although there is evidence of a correlation in weather and the day of the week (Cerveny and Balling 1998), which we were surprised to discover in our data, the correlations are tiny, so that including a set of day-of-the-week fixed effects does essentially nothing to change the results. Individual's true health is also suspected to be a confounding factor here, because one's health clearly may influence participation in outdoor recreational activities, but may also be influenced by the weather, such as through seasonal allergies. Since the identification is off intra-month weather fluctuations at a particular time of the day, we think it is unlikely that our results will reflect heath effects of the weather. Nonetheless, we also check the robustness of our estimates to this possibility by directly controlling for one's health status in the analysis, using a measure of self-reported health available in the GSS. As it turns out, this also does essentially nothing to change the estimates.

With our estimates of the parameters using equation(2.4.1) and equation (2.4.2), we

[4]The regression analysis of this specification is conducted using a probit model, which will be estimated with a maximum likelihood procedure.

can predict a probability of an average person engaging in an average outdoor activity, given any combination of weather conditions (whether in sample or not). To check the meaningfulness of this weather index, we conclude by using regression analysis to examine the robustness of our results and, in particular, the functional form we employ. Compared to equation (2.4.2), which is the specification based on the theory-oriented approach, the stepwise approach is a more data-driven approach. In the stepwise approach, the regression analysis begins with a very general specification that allows thermal and aesthetic effects even under physical conditions, in which the linear index v_{ict} in equation (2.4.1) is given by:

$$\begin{aligned} v_{ict} = \mathbf{p_{ct}}[&\alpha_{1p}h_{ct} + \alpha_{2p}h_{ct}^2 + \alpha_{3p}h_{ct}^3 + \alpha_{4p}h_{ct}^4 + \alpha_{5p}w_{ct} + \alpha_{6p}w_{ct}^2 + \\ &\alpha_{7p}w_{ct}^3 + \alpha_{8p}w_{ct}^4 + \alpha_{9p}(h_{ct} * w_{ict}) + \alpha_{10p}(h_{ct}^2 * w_{ct}) + \alpha_{11p}d_{ct} + \\ &\mathbf{z}_c\gamma_p + \mathbf{x}_t\delta_p] + (\mathbf{1} - \mathbf{p_{ct}})[\alpha_{1np}h_{ct} + \alpha_{2np}h_{ct}^2 + \alpha_{3np}h_{ct}^3 + \alpha_{4np}h_{ct}^4 + \\ &\alpha_{5np}w_{ct} + \alpha_{6np}w_{ct}^2 + \alpha_{7np}w_{ct}^3 + \alpha_{8np}w_{ct}^4 + \alpha_{9np}(h_{ct} * w_{ict}) + \\ &\alpha_{10np}(h_{ct}^2 * w_{ct}) + \alpha_{11np}d_{ct} + \mathbf{z}_c\gamma_{np} + \mathbf{x}_t\delta_{np}] \end{aligned} \tag{2.4.3}$$

The effects of humidex and wind are estimated with a quartic function form, which intends to capture a more complicated nonlinear effects these factors can potentially have on outdoor leisure activities. A same set of control variables as in equation (2.4.2), such as city where the individual lives and month of the year, are also included in the analysis. The estimation, which is also estimating the probit model with maximum likelihood method, takes an approach called backward elimination. In the backward elimination process, the estimation starts with all candidate variables listed in equation (2.4.3), tests the deletion of each variable using the chosen criterion, which is the significance level indicated by p-values, deletes the variable (if any) that improves the model the most by being deleted, and repeats this process until no further improvement is possible. Specifically, the weather terms are dropped from the specification if the significance level of the term is lower than 10%. Therefore, the final function form of the stepwise analysis only contains weather terms, of which the effects on outdoor activities are statistically significant at the level of 10% or higher. Since the specification of equation (2.4.2) is nested in this equation that the step-wise analysis starts with, we could use the final function form from the stepwise analysis to check the robustness of the function form employed for constructing the weather index in equation (2.4.2).

2.5 Results

The results from the nonparametric analysis are presented in Figure 2.2. The likelihood of people engaging in outdoor leisure activities is always higher when there are no physical weather conditions (no rain or strong wind), and the more pronounced correlation between humidex and outdoor activities in the absence of physical weather conditions is consistent with the notion of an overriding effect of physical weather conditions. It is interesting to notice that in the graph, the likelihood of people participating in outdoor leisure activities peaks at a humidex of 27 ℃. Mieczkowski(1985) says, "In more general terms, researchers agree that the optimum of comfort for a lightly dressed, seated person lies around 20 ℃ – 27 ℃ DBT and between 30 and 70 percent RH. (p.223)"[5] These optima values of dry bulb temperatures and relative humidity Mieczkowski(1985) states correspond to humidex values between 20 ℃ and 33 ℃, which implies the peak point 27 ℃ falls into the range of the most comfortable humidex. Moreover, the humidex does appear to promote outdoor activities in the lower range but deter those activities in the higher range, which supports our speculation of the nonlinear effect of humidex in (2.4.2). The effect of wind also appears to be nonlinear and peaks around 17km/hr. In comparison to humidex, the nonlinear patterns of the effects of wind and cloud on outdoor activities are less pronounced.

Table 2.2 reports the main results of estimating (2.4.2). Consistent with the findings from the nonparametric analysis, warmer weather results in a higher propensity of engaging in outdoor recreation up to some threshold humidex, the value of which depends on the amount of wind. In the situation when there is no wind, a one standard deviation increase in the humidex above the 25th percentile increases the probability of engaging in outdoor leisure activities from 0.0599 to 0.0607. In contrast, at the 75th percentile of the humidex a standard deviation increase of humidex decreases the probability of outdoor recreation from 0.1295 to 0.1238.

Wind primarily has the effect of flattening the humidex function, so that increases in the humidex, whether they lie above or below the threshold, have smaller marginal effects. For example, when there is no wind, an increase in the humidex from 10 ℃ to 15 ℃ increases the probability of outdoor recreation from 0.0969 to 0.1295, which is a 33.6% increase. In comparison, the same increase in the humidex would lead to only a 23.3% (from 0.1074 to 0.1324) increase in outdoor recreation if wind speed is at 15km/hr. This is because at low temperatures an increase in the humidex increases outdoor recreation, but the wind dampens the effect of this increase in the humidex on people's thermal sensation. The opposite is true when the humidex is at the upper end of the distribution. When the humidex is 30 ℃ and increases by 5 degrees, the wind dampens the negative effect of the increase in the humidex. Specifically, the probability of outdoor recreation decreases from 0.1640 to 0.1557 if the wind speed is 15km/hr, compared with a decrease from 0.1295 to 0.0970 (25.09% decrease) when there is no wind. Over the entire estimated function, the "bliss point" combination of weather conditions is a humidex of 27.2 ℃; a wind speed of 14.7 km/hr; and clear skies. Also, over virtually our entire sample, physical conditions (rain and high wind speed), which negate the effects of the humidex, wind speed and cloud cover, result in less outdoor recreation.

[5] DBT and RH denote dry bulb temperature and relative humidity, respectively.

Having estimated (2.4.2), we can go back to our weather data and for every city-day-hour weather observations, predict a probability of being outdoors or a "weather quality". To demonstrate the meaningfulness of this index, in Figure 2.3 we plot it using average daily weather conditions between 1976 and 2008 from six Canadian cities – Vancouver, Edmonton, Winnipeg, Toronto, Montreal, and Saint John's. These six large cities are chosen because they capture the main climate differences across Canada's major urban centres. Since the predicted values are based on a common city-month-hour reference group (Toronto, July, and 2pm), the variations purely reflect differences in weather conditions, as opposed to variations in individual outdoor recreation preferences across cities or time. The results are entirely consistent with popular perceptions. Vancouver enjoys better Spring weather, but summers in Toronto and Montreal are preferred. Averaging the city predicted weather quality index from April to November, the integrated city weather profiles imply that Toronto enjoys the highest average weather quality, followed by Vancouver, Montreal, Winnipeg, Edmonton, and St. John's.

To provide us with more assurance of our chosen theory-driven specification, in Table 2.3 we report the results of a more data-driven approach – stepwise regression analysis. The overriding effect of physical conditions the theoretical approach emphasizes is confirmed by the fact that most weather terms are dropped due to their low level of significance when physical weather conditions are present. Cloud cover appears to be the main influential factor under physical weather conditions, and it decreases the the value of the index. In the absence of physical weather conditions, the effect of humidex, wind speed, and cloud cover all appear influential. Although the quadratic humidex term is not significant, instead, the cubic term is. However, when we include the cubic term in equation (2.4.2) together with the quadratic term, its significance disappears and the magnitude is small. Moreover, holding other variables constant, the linear and cubic effects of humidex identified in the stepwise regression and the linear and quadratic effects of humidex estimated in the theory-oriented approach have a correlation of 0.98, suggesting that whether we include a quadratic or cubic term makes little difference. The stepwise regression identifies linear and quadratic effects of the wind, which is consistent with the terms included in the theory-oriented approach in equation (2.4.2). Moreover, the magnitudes of the estimates are close. Overall, the estimates of the stepwise approach produces very similar rankings of weather conditions as the theory-oriented approach does, as where the weather indices have a correlation of 0.95 in our sample of weather data. This justifies our use of the theoretical approach to identify the functional form of the weather index.

A complication in our approach is that the set of 13 "clearly outdoor" activities we focus on almost certainly misses many outdoor activities. For example, many of the swimming episodes in the data are presumably outdoors. As a result, weather fluctuations that lead individuals to substitute from an indoor activity to swimming outside will be missed and the estimates will be attenuated. To examine the robustness of the estimated index to this potential bias, we have tried estimating equation (2.4.2) dropping "ambiguous" leisure activities, such as swimming. As the results reported in Table 2.4 indicate, the main effect of doing this is to shift up the intercept rather than change the shape of the function. For example, the optimal humidex and wind speed are 26.8 ℃ and 14.5 km/hr, respectively, when we exclude all ambiguous leisure activities. When we go one step further excluding all home production activities, many of which may also be outdoors, the results also remain similar, where the optimal values are 27.3 ℃ for the humidex and 16.4km/hr for the wind speed.

The weather index is constructed with the primary objective to quantitatively measure weather quality in order to interpret the correlation between absenteeism and weather quality as reflecting employees shirking behaviour in Chapter 3. To motivate this analysis, a preliminary labour supply analysis is presented in this chapter using GSS time-use data. The sample for this analysis consists of the same types of individuals used to estimate weather index – full-time wage or salary workers with regular day-time shifts. However, in this analysis we restrict the observations to activities done between 9am to 5pm from Monday to Friday, since these are the time that these workers are expected to face contractual work hour obligations. Using the information on detailed activities, we group activities into 3 categories: "present at work"; "participating in outdoor leisure activities"; "participating in other activities".

Using a multinomial logit regression, the effect of weather quality on the relative probabilities of workers being absent from work is presented in Table 2.5. As weather quality improves up to an index value of 0.21, the 95th percentile of the weather index distribution in our sample, workers appear to be increasingly likely to be absent from work and participating in outdoor leisure activities. In other words, weather quality does appear to be positively correlated with observed work absence. Of course, this absence-weather correlation does not necessarily reflect malfeasant shirking behaviour. Nonetheless, this finding suggests that analysis using data that contains more detailed information on employees' labour market behaviour, such as the reasons of being absent from work, is a worthwhile endeavour. Exploring the absence-weather correlation can provide valuable insights in understanding employees' behaviour, such as absenteeism and shirking, as well as policy implication for employers to effectively manage workplace absence, which is the main objective of the following chapter.

2.6 Summary

In contrast to currently-available climate indices, this chapter constructs an index to represent people's weather preference over a broader set of outdoor leisure activities. Linking hourly weather data with GSS time-use data, in which respondents' detailed activities are recorded, we construct a weather quality index by predicting the probability of individuals' voluntary participation in outdoor high-utility leisure activities conditional on thermal, aesthetic and physical weather conditions.

The results show that physical weather conditions – rain and high wind speeds – are highly influential in deterring people's outdoor activities. In the presence of these physical conditions, not only is the weather quality index significantly lower than otherwise, but the effects of other weather elements in determining the weather quality are nullified. In contrast, when there is no physical weather condition, humidex plays an important role in determining weather quality. In particular, the humidex appears to promote outdoor recreation at lower values and become an obstacle when it becomes too high. The optimal value of the humidex is determined jointly with wind speed, which is around 27 ℃ and 14 km/hr, respectively. The meaningfulness of our weather quality index is tested by comparing the predicted weather quality across 6 most representative Canadian cities and the results are consistent with popular perceptions.

The weather index in this chapter is constructed to quantitatively evaluate weather quality by the probability to engage in outdoor recreational activities under a particular type of weather conditions. This weather index not only provides a ranking of different weather conditions in terms of their values to outdoor leisure, but also reflects the weather-induced variations in the utility of the these outdoor activities. The theoretical foundation of using weather to study time allocation decisions is that weather fluctuations can change relative utility gain across different activities mainly thorough its significant effect on outdoor activities. Therefore, this weather index directly captures the underlying force that drives substitutions between activities. For studying time allocation decisions, such as time substitution between work and leisure, this transparent interpretation of the index values is greatly helpful for understanding causal mechanism behind the changes in time allocations.

Table 2.1: Enjoyability of respondents' daily activities

Activities	Number of observations	Percentage	Probability most enjoyed
Outdoor leisure	2,066	6.09%	40.83%
Indoor leisure	6,600	19.45%	16.09%
Ambiguous leisure	8,018	23.65%	21.66%
Domestic work	5,323	15.70%	4.67%
Care giving for household members	1,512	4.46%	20.75%
Shopping and services	2,599	7.66%	5.99%
Personal care	6,260	18.46%	6.55%
School and education	267	0.79%	15.87%
Organizational, voluntary and religious activity	1,263	3.72%	16.75%
Total	33,908	100.00%	

Notes: 1. *Outdoor leisure:* (activities taken place outdoors): gardening; walking, hiking, jogging, and running; golf; fairs, festivals, circuses, parades; bicycling; pleasure drives; fishing; boating; rowing, canoeing, kayaking, and wind surfing and sailing; zoos; camping; hunting; horseback riding, rodeo, jumping, and dressage. *Indoor leisure:* (activities taken place indoors): watching television; computer use; exercise, yoga, weight lifting; telephone conversation; video games, computer games; games, cards, puzzle, board games; movies, films; listening to CD's, cassette tapes or records; casino, bingo, arcade; bowling, pool, ping-pong, pinball; opera, ballet, and theater; museums; judo, boxing, wrestling, and fencing; art galleries. *Ambiguous leisure:* (activities taken place either outdoors or indoors): socializing with friends/relatives; relaxing, thinking, resting, and smoking; reading books; talking, conversation, and telephone; reading newspapers; socializing at bars, clubs; other social gathering; amateur sports; swimming and waterskiing; football, basketball, baseball, volleyball, hockey, soccer, and field hockey; domestic home crafts done for pleasure; professional sports event; music, theatre, and Dance; listening to radio; coaching; tennis, squash, racquetball, and paddleball; skiing, ice skating, sledding, curling, and snowboarding; heritage site; pop music and concerts.
Source: Activity data from 1992, 1998 and 2005 General Social Survey (GSS).

Table 2.2: Theoretical specification probit estimates of the effect of weather on the incidence of outdoor recreational activity

	Coefficient	Standard error
Physical conditions	0.3951*	0.2203
Humidex	0.0767***	0.0212
Humidex^2/100	−0.1704***	0.0508
Wind	0.0252*	0.0130
Wind^2/100	−0.0417**	0.0197
Humidex*Wind/100	−0.2106*	0.1098
Humidex^2*Wind/1000	0.0603**	0.0267
Cloud	−0.0167***	0.0053
April	−0.2089**	0.0916
May	0.0612	0.0730
June	−0.0430	0.0647
August	−0.0536	0.0667
September	−0.1175	0.0727
October	−0.1970**	0.0810
9 a.m.	−0.2794***	0.0662
10 a.m.	−0.1552***	0.0595
11 a.m.	−0.0580	0.0514
12 p.m.	−0.1577***	0.0493
1 p.m.	−0.0425	0.0385
3 p.m.	−0.0255	0.0390
4 p.m.	−0.1484***	0.0487
5 p.m.	−0.4080***	0.0581
6 p.m.	−0.6214***	0.0524
7 p.m.	−0.5197***	0.0519
8 p.m.	−0.5443***	0.0532
9 p.m.	−0.8112***	0.0597
City fixed effects	yes	–
Constant	−1.8960***	0.2452
Pseudo R^2	0.0662	
N	33,834	
Optimal humidex	27.2 ℃	
Optimal wind speed	14.7 km/hr	

Notes: Standard errors are clustered by city and time (month, day, and hour). 74 observations are dropped as they predict failure perfectly. ***,** ,* indicate statistical significance at the 1%, 5%, and 10% levels, respectively.
Source: Activity data from 1992, 1998 and 2005 General Social Survey (GSS). Weather data from Canadian National Climate Data and Information Archive (NCDIA).

Table 2.3: Stepwise probit estimates of the effect of weather on the incidence of outdoor recreational activity

	Coefficient	Standard error
No physical conditions*Humidex	0.0456***	0.0116
No physical conditions*Humidex^2/100	0.0000	–
No physical conditions*Humidex^3/1000	−0.0267***	0.0080
No physical conditions*Humidex^4/10000	0.0000	–
No physical conditions*Wind	0.0222*	0.0115
No physical conditions*Wind^2/100	−0.0412**	0.0194
No physical conditions*Wind^3/1000	0.0000	–
No physical conditions*Wind^4/10000	0.0000	–
No physical conditions*Humidex*Wind/100	−0.1820*	0.0986
No physical conditions*Humidex^2*Wind/1000	0.0541**	0.0250
No physical conditions*Cloud	0.0000	–
No physical conditions*Cloud^2/100	0.0000	–
No physical conditions*Cloud^3/1000	−0.1628***	0.0500
No physical conditions*Cloud^4/10000	0.0000	–
Physical conditions*Humidex	0.0000	–
Physical conditions*Humidex^2/100	0.0000	–
Physical conditions*Humidex^3/1000	0.1134**	0.0550
Physical conditions*Humidex^4/10000	0.0000	–
Physical conditions*Wind	0.0000	–
Physical conditions*Wind^2/100	0.0000	–
Physical conditions*Wind^3/1000	−0.0032*	0.0017
Physical conditions*Wind^4/10000	0.0000	–
Physical conditions*Humidex*Wind/100	0.0000	–
Physical conditions*Humidex^2*Wind/1000	0.0000	–
Physical conditions*Cloud	0.2950***	0.0871
Physical conditions*Cloud^2/100	−0.0280***	0.0081
Physical conditions*Cloud^3/1000	0.0000	–
Physical conditions*Cloud^4/10000	0.0000	–
Month fixed effects	yes	–
Hour of the day fixed effects	yes	–
City fixed effects	yes	–
Constant	−2.0633***	0.1837
Pseudo R^2	0.0679	
N	33,834	
Optimal humidex	27.2 ℃	
Optimal wind speed	14.7 km/hr	

Notes: Standard errors are clustered by city and time (month, day, and hour). Regression also controls for city, month and hour. 74 observations are dropped as they predict failure perfectly. ***,** ,* indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Weather terms are dropped when significance is lower than 10%.
Source: Activity data from 1992, 1998 and 2005 General Social Survey (GSS). Weather data from Canadian National Climate Data and Information Archive (NCDIA).

Table 2.4: Robustness analysis probit estimates of the effect of weather on the incidence of outdoor recreational activity (excluding potential ambiguous leisure and home production activities)

	Outdoors (excl. ambiguous leisure activities)		Outdoors (excl. ambiguous leisure activities and home production)	
	Coefficient	Standard error	Coefficient	Standard error
Physical conditions	0.3884*	0.2310	0.3051	0.2895
Humidex	0.0797***	0.0223	0.0882***	0.0290
Humidex^2/100	-0.1736***	0.0537	-0.1910***	0.0720
Wind	0.0249*	0.0136	0.0215	0.0186
Wind^2/100	-0.0413**	0.0207	-0.0555*	0.0287
Humidex*Wind/100	-0.2138*	0.1152	-0.1703	0.1643
Humidex^2*Wind/1000	0.0618**	0.0280	0.0579	0.0404
Cloud	-0.0188***	0.0056	-0.0192***	0.0073
April	-0.2791***	0.0966	-0.4540***	0.1271
May	0.0093	0.0772	-0.0048	0.1031
June	-0.0742	0.0694	-0.1062	0.0958
August	-0.0901	0.0713	-0.2029**	0.0920
September	-0.1759**	0.0771	-0.3511***	0.0970
October	-0.2476***	0.0860	-0.3514***	0.1123
9 a.m.	-0.3790***	0.0690	0.1600	0.1252
10 a.m.	-0.2419***	0.0626	-0.0255	0.1063
11 a.m.	-0.1227**	0.0546	0.0316	0.0990
12 p.m.	-0.2167***	0.0523	0.0034	0.0931
1 p.m.	-0.0575	0.0413	0.0319	0.0684
3 p.m.	-0.0237	0.0422	-0.2032***	0.0676
4 p.m.	-0.1051**	0.0527	-0.2888***	0.0844
5 p.m.	-0.3772***	0.0632	-0.5976***	0.0987
6 p.m.	-0.6926***	0.0561	-1.0106***	0.0888
7 p.m.	-0.5647***	0.0555	-1.1236***	0.0860
8 p.m.	-0.5874***	0.0567	-1.3992***	0.0851
9 p.m.	-0.8717***	0.0632	-1.8304***	0.0908
City fixed effects	yes	–	yes	–
Constant	-1.7106***	0.2577	-0.5125	0.3233
Pseudo R^2	0.0760		0.2382	
N	25,836		8,655	
Optimal humidex	26.8 ℃		27.3 ℃	
Optimal wind speed	14.5 km/hr		16.4 km/hr	

Notes: Standard errors are clustered by city and time (month, day, and hour). ***,** ,* indicate statistical significance at the 1%, 5%, and 10% levels, respectively.
Source: Activity data from 1992, 1998 and 2005 General Social Survey (GSS). Weather data from Canadian National Climate Data and Information Archive (NCDIA).

Table 2.5: Multinomial logit estimates of the probability of absence from work

	Outdoor leisure activities		Other activities	
	Coefficient	Standard error	Coefficient	Standard error
Weather	24.4180***	8.0194	0.0567	3.2012
Weather^2	-60.5925**	29.5764	-5.7766	12.3482
Job acquisition rate	1.8390	1.6109	-0.8953	0.6479
Union	0.6540**	0.2688	0.4169***	0.1006
April	-0.9889*	0.5251	-0.6089***	0.1777
May	-0.4414	0.3082	-0.3246**	0.1486
June	-0.9492***	0.3487	-0.4233***	0.1419
August	-0.6519*	0.3422	-0.2054	0.1523
September	-0.8741**	0.3877	-0.7138***	0.1486
October	-0.7774	0.5836	-0.3346*	0.1752
Tuesday	-0.6173*	0.3206	-0.3872***	0.1229
Wednesday	-0.1647	0.3043	-0.3451***	0.1230
Thursday	-0.3390	0.3183	-0.3071**	0.1304
Friday	0.2140	0.3292	-0.0029	0.1219
9 a.m.	0.2111	0.3184	0.0297	0.0857
10 a.m.	0.0135	0.2480	-0.1390**	0.0619
11 a.m.	0.0109	0.2002	-0.2370***	0.0445
12 p.m.	0.4025*	0.2242	0.3463***	0.0614
1 p.m.	0.4360***	0.1566	0.2122***	0.0416
3 p.m.	0.1292	0.1296	0.0607*	0.0337
4 p.m.	0.7266***	0.1927	0.6016***	0.0581
5 p.m.	1.6849***	0.3020	1.7246***	0.1035
City fixed effects	yes	–	yes	–
Constant	-7.3038***	1.5398	-1.5200***	0.5494
Pseudo R^2		0.1024		
N		23,056		

Notes: Standard errors are clustered by city and time (month, day, and hour). ***,**,* indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The excluded group in the multinomial logit analysis is workers being at work between 9a.m. to 5p.m. from Monday to Friday. Other activities include indoor leisure activities, ambiguous leisure activities, domestic work, care giving for household members, shopping and services, personal care, school and education, and organizational, voluntary and religious activities.
Source: Activity data from 1992, 1998 and 2005 General Social Survey (GSS). Weather data from Canadian National Climate Data and Information Archive (NCDIA).

Figure 2.1: Sample distribution of weather variables

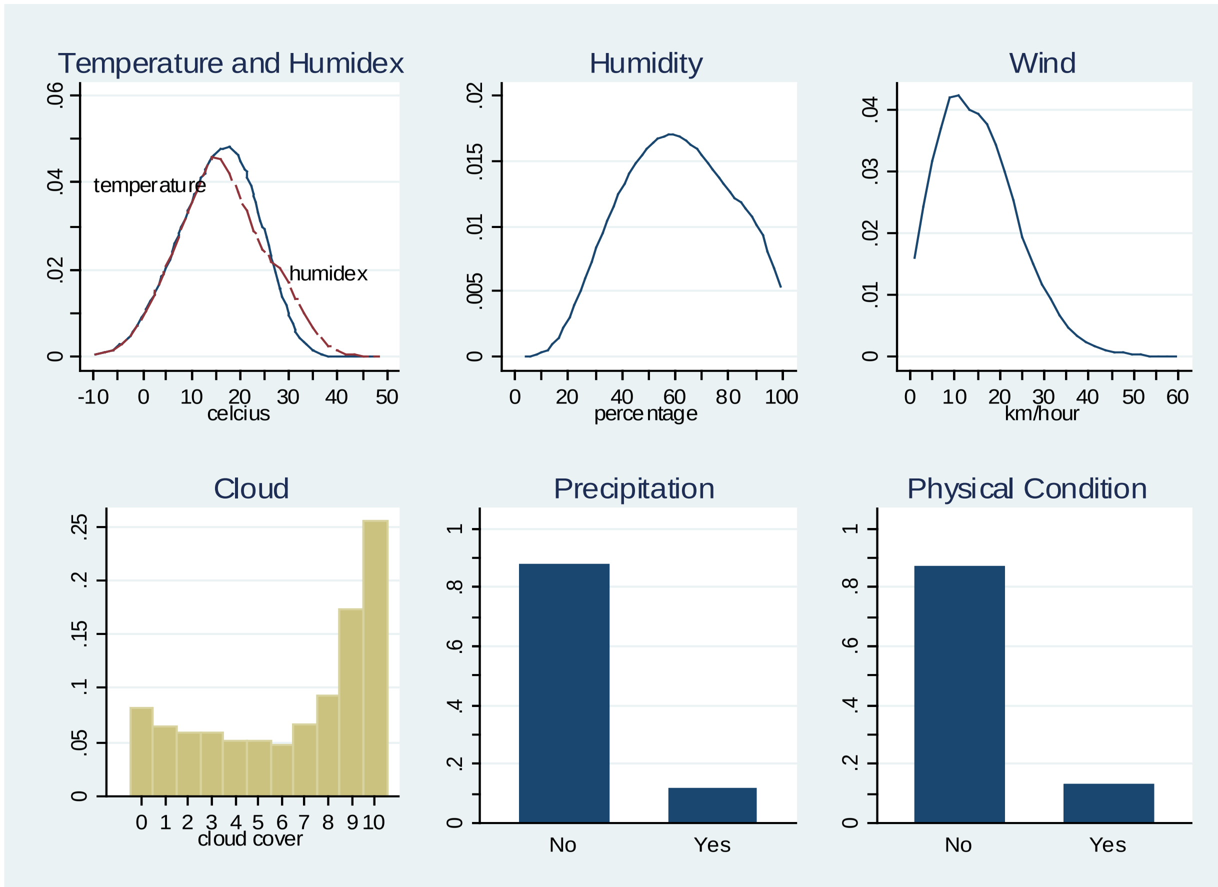
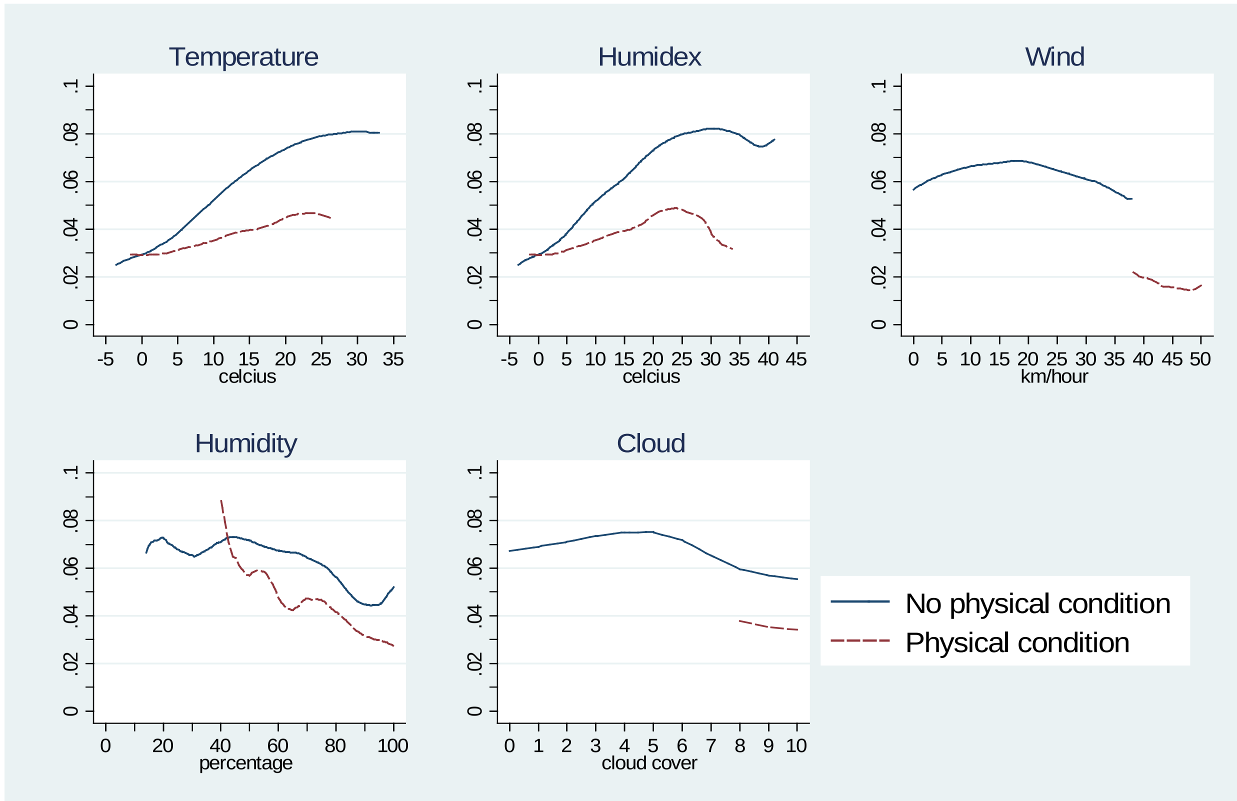
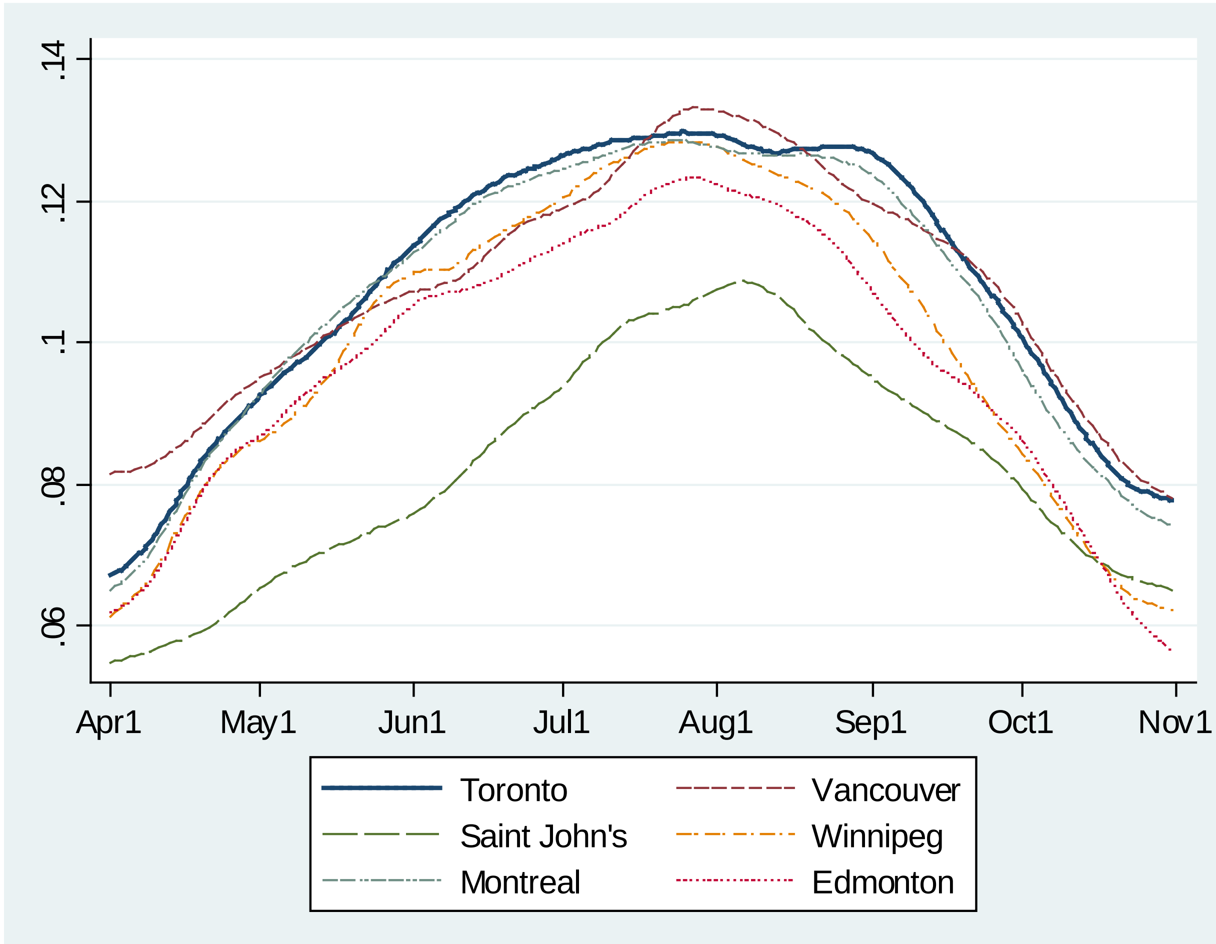


Figure 2.2: Nonparametric analysis of probability of outdoor recreation and weather



Notes: Vertical axis plots the predicted probability of outdoor recreation from using kernel density nonparametric regression analysis.

Figure 2.3: April to October weather quality in six Canadian cities



Notes: Vertical axis plots the predicted probability of outdoor recreation from (2.4.2) using average daily (9am-9pm) weather conditions between 1976 and 2008. All values are predicted for the reference group (Toronto at 2pm in July), so that all variation purely reflects weather variations, and not variation in weather preferences across cities or time.

Chapter 3

Gone Fishing! Reported Sickness Absenteeism and the Weather

3.1 Introduction

On the surface, labour contracts between employers and employees are straightforward. An employer pays a worker a predetermined wage for providing a level of effort in the execution of some task that is of value to the employer. As long as the employee's effort level is easily observed or monitored, such as in situations where piece rates are paid, the contract is complete and no complications arise. If, however, effort levels are imperfectly monitored, a principal-agent problem arises, in which workers have an incentive to shirk contractual work obligations.

In situations where employee work effort is difficult to monitor, theories across disciplines have suggested that changing the shirking incentives that employees face can reduce employees' shirking activities. Outside economics, much of the focus has been on the monitoring technology itself. Economists, in contrast, have tended to focus more on the expected cost of shirking activity, and in particular the cost associated with dismissal in the event that shirking activity is detected. For example, in explaining persistent wage rigidities and involuntary unemployment in competitive labour markets, the efficiency wage model of Shapiro and Stiglitz (1984) suggests that employers may be willing to pay higher than market-clearing wage rates in order to increase the wage loss that employees experience when dismissed. However, before employers and policy makers can use theories such as this to inform optimal policy design, it is critical that the real-world relevance of these models be evaluated.

The key challenge in empirical studies of employee shirking behaviour is the measurement of shirking activity, which is by its very nature difficult to identify. It is, of course, the inherent unobservability of shirking that provides workers with an incentive to shirk. Within the economics literature, the most common approach to measure shirking activity has been to focus on differences in levels of sickness absenteeism either over time or across workers or workplaces. However, given that to much, and perhaps the vast majority, of sickness absenteeism reflects individuals' genuine health status, using overall sickness absenteeism to understand shirking incentives is problematic.

Chapter 1 proposes a model of sickness absenteeism in which the workers' marginal utility of outdoor leisure is increasing in the quality of the weather. Moreover, this effect of the weather is larger when workers are healthier. As a consequence of these preferences, the model implies that although much of overall reported sickness absenteeism may be entirely legitimate, any observed correlation between the weather and reported sickness necessarily reflects the behaviour or the inframarginal workers who are the most healthy, and is therefore malfeasant in nature. To the extent that the underlying assumptions of this theoretical model are plausible, this suggests that one can empirically relate data on levels of reported sickness absenteeism to an index of weather quality to identify *marginal* changes in shirking activity. With this measure of shirking activity, the model provides two additional propositions that can be used to further our understanding of how various shirking incentives, such as the level of sick pay or the probability of replacing a lost job, influence shirking activities. First, the model implies that levels of sickness absenteeism are less responsive to weather improvements if workers are already facing high shirking incentives, such as where the generosity of sick pay is high or the likelihood of finding a replacement job is high. Second, if employers pay efficiency wages to try to augment shirking incentives, these interaction effect of the weather and other shirking incentives is attenuated, and in the case of sick pay reversed, providing us a strategy for obtaining some indirect evidence of the real-world relevance of the efficiency wage hypothesis.

Linking 12 years of employee data from Canada's monthly Labour Force Survey (LFS), which unlike the Current Population Survey (CPS) in the U.S., regularly queries reasons for short-term absences, with archival weather data from 56 Canadian cities, we find clear evidence of a positive relation in the non-winter months between the quality of outside weather conditions and reported sickness absenteeism among employees working indoors. Examining this absenteeism-weather relation across workers facing different existing shirking incentives, such as between salaried workers, who are typically paid for time off work, and hourly-paid workers, who are not, the results are also supportive of the second proposition. That is, we almost consistently estimate larger marginal effects of the weather among workers where existing shirking incentives are the lowest. Moreover, this appears to be true regardless of whether or not we condition on an estimated individual wage premium. Given the third proposition of Chapter 1, these empirical results offer little evidence to support the mechanisms underlying the efficiency wage hypothesis.

The main limitation of our empirical strategy is that we are only able to identify marginal changes, as opposed to levels, in one particular type of shirking behaviour – shirking contractual work hours to take advantage of good weather conditions. However, we believe our approach provides a better approach for identifying and furthering our understanding of employee shirking behaviour in two important respects. First, once we condition on where an individual lives and the time of the year, weather conditions are unambiguously exogenous. That is, sample variations in the weather conditions facing employees are necessarily unrelated to the unobservable attributes of employees. Consequently, we can confidently interpret the weather-absenteeism relation as direct causal effect of the weather, which given our focus is indoor workers employed in the non-winter months, we argue the weather-absenteeism behaviour is more likely to reflect the influence of the weather on the marginal value of outdoor leisure, as opposed to a direct effect of the weather on employee health.[1] Second, the absenteeism-weather relation this study focuses on arguably better captures employees' shirking behaviour than what has been used elsewhere. Most notably, unlike the employee dismissal rates examined by Cappelli and Chauvin (1991), which we take to be the most credible attempt to directly measure shirking in the current literature, our measure potentially captures not only shirking that is detected, but also that which goes undetected.

The following section of the chapter provides a more complete review of the existing empirical literature that uses absenteeism as the measure for shirking and examines the potential mechanisms explaining weather-absenteeism correlations. The existing literature studying the response of shirking behaviour to shirking incentives, in particular the incentives offered by efficiency wages will also be reviewed in section 2. In section 3, the empirical methodology and data used to test the propositions in Chapter 1 are introduced. Section 4 discusses the results and provides a number of robustness checks of the results, followed by a concluding section, which summarizes.

3.2 Related Literature

This study identifies marginal changes in shirking activity by exploiting the link between the weather and reported sickness absenteeism. As noted above, there is a strong precedent within the economics literature for using data on sickness absenteeism to study employee shirking behaviour. In fact, sickness absenteeism appears to be the most popular measure of shirking in empirical studies. Leigh (1985), Audus and Goddard (2011), Arai and Thoursie (2005), and Askildsen, Bratberg and Nilsen (2005) interpret procyclical rates of sickness absenteeism as consistent with an increased cost of shirking during recessions,

[1] Note that the theoretical model in Chapter 1 avoids this complicating alternative interpretation of the weather-absenteeism relation by assuming that the distributions of the weather (λ) and health (θ) state-dependent variables, are independent.

when lost jobs are less easily replaced. Using data from a large Italian bank, Ichino and Riphahn (2005) identify a sharp discontinuous rise in employee absenteeism precisely when probationary periods end and legislative protections against firing kick in. Frick and Malo (2008) interpret variation in absenteeism rates across EU countries as resulting primarily from different sickness benefits. And Bradley, Green and Leeves (2007) interpret changes in absenteeism among Australian teachers who change schools as evidence of how workplace absence norms affect shirking incentives. The identification in these papers is, however, complicated by the possibility that observed variations in sickness absenteeism reflect genuine health. Procyclical sickness absenteeism is consistent with the cost of dismissal influencing shirking incentives, but also with evidence of countercyclical health (Ruhm, 2000; Charles and DeCicca, 2008). Group-interaction effects are consistent with local or workplace absence norms influencing shirking incentives, but also with contagious illnesses. And sick pay provision, whether from a company or legislative, is ultimately endogenous, potentially driven by the preferences of workers or voters with varying health. The only empirical papers in the economics literature to consider an independent role of health in determining absenteeism are those concerned with gender differences (Paringer, 1983; Vistnes, 1997; and Ichino and Moretti, 2006). And interestingly, these papers consistently find that health factors are, if anything, more important in explaining observed absenteeism than economic factors that might affect shirking incentives. Therefore, using overall sickness absenteeism to understand shirking decision is problematic.

The analysis in this chapter is not the first to study the empirical link between absenteeism and the weather. Searching across disciplines there are three papers, which are all from outside economics, that examine the absenteeism-weather link: one from psychology (Mueser, 1953); one from epidemiology (Pocock, 1972); and one from environmental science (Markham and Markham, 2005). In all cases, the correlation is interpreted as either the effect of weather on physical or mental health (e.g., influence of humidity on arthritic pain) or on the cost of getting to work (e.g., during a snow storm). The possibility of weather influencing shirking incentives is only ever mentioned as an afterthought. In the discussion of his results Mueser, for example, writes: "It is easy to imagine that when it was sunny and beautiful outside the chore of earning a livelihood was put off."(p.337) Interestingly, however, the implied correlation arising from shirking is opposite in sign from that working through health or travel costs, if we believe that higher weather quality is associated with improved health and lower travel costs. Evidence of a positive weather-absenteeism correlation would, therefore, suggest that even if these alternative mechanisms are at play, the influence of the weather on shirking incentives dominates.

We are also not the first in the economics literature to exploit the exogeneity of the weather. Roll (1984) and Angrist, Graddy and Imbens (2000) use weather conditions at sea and in Central Florida to identify demand functions for fish and oranges, respectively. Paxson (1992) uses regional rainfall in Thailand to construct estimates of shocks to transitory income of Thailand farm households, which are then used to study these household's saving behaviour. Boustan, Fishback and Kantor (2010) use extreme weather events to instrument migrant flows, while Burke and Leigh (2010) use adverse weather shocks to instrument output contractions. A much larger literature, however, is the research relating the weather to stock market returns and trading activity, which has been interpreted as evidence of the psychological effects of the weather, specifically tastes for risk, and against the efficient markets hypothesis of fully rational price setting (e.g., Saunders (1993); Hirshleifer and Shumway (2003); Jacobsen and Marquering (2008)). Lastly, Connolly (2008) relates daily hours of work to the incidence of rain and finds evidence of intertemporal labor supply responses to weather. But with no information in her data on the nature of the hours adjustments, in particular whether they reflect sickness absenteeism, overtime, vacation days, or even private time-in-lieu-of arrangements with employers, her results do not necessarily tell us anything about shirking activity.[2]

The efficiency wage model of Shapiro and Stiglitz (1984) is unquestionably the most important and well-recognized theoretical model in the economics literature that speaks to the phenomenon of employee shirking. The appeal of the theory, and why it continues to be included as part of the standard curricula in upper-year macroeconomics courses, is that it offers an explanation of both equilibrium wage dispersion and involuntary unemployment in a world of perfectly homogeneous workers. However, the power of this explanation rests on the underlying mechanism that the wage rates can be used as an instrument to augment employee shirking incentives. There now exists a considerable empirical literature attempting to test the real-world relevance of the this efficiency wage hypothesis. Although the approaches to testing this hypothesis are quite varied, these studies share the underlying limitation that they lack of a direct measure of shirking activity. For example, in one of the most frequently cited empirical studies of efficiency wages, Krueger and Summers' (1988) found a negative relationship between wage differentials and employee turnover, and this finding has been seen to be indicative of the effectiveness of wage premia implied by the theory. However, in their critical review, Murphy and Topel (1990) are skeptical arguing that systematic wage differences across industries are entirely consistent, both theoretically and empirically, with the sorting of workers on unobservable dimensions. For example, it may simply be that workers who tend to be more loyal, and therefore less likely to quit their jobs, also tend to be more productive, and therefore earn higher wages. Or

[2]Searching across disciplines we have found three papers outside economics that examine the absenteeism-weather link – one from psychology (Mueser (1953)); one from epidemiology (Pocock (1972)); and one from environmental science (Markham and Markham (2005)). In all cases, the correlation is interpreted as either the effect of weather on physical or mental health (e.g., influence of humidity on arthritic pain) or on the cost of getting to work (e.g., during a snow storm). The possibility of weather influencing shirking incentives is only ever mentioned as an afterthought. For example, in the discussion of his results, Mueser writes: "It is easy to imagine that when it was sunny and beautiful outside the chore of earning a livelihood was put off."

perhaps higher loyalty actually serves to increase wage outcomes of employees directly. In conclusion, Murphy and Topel (1990) call for more direct tests relating wage differentials to characteristics of industries that provide incentives for employers to pay wage premia. Following on this suggestion, Neal (1993) relates inter-industry wage differentials to a measure of the frequency to which employees' work is monitored contained in the U.S. Panel Study of Income Dynamics (PSID). Chen and Edin (2002), on the other hand, compare inter-industry wage differentials between Swedish production workers paid piece rates as opposed to wage rates based on time. While the results of Chen and Edin are overwhelmingly mixed, Neal's findings tend to contradict the predictions of the shirking model of efficiency wages.

Using firm-level productivity data, as in Huang et al. (1998) and Wadhwani and Wall (1991) respectively, more direct evidence of the effect of the efficiency wage hypothesis is provided, but still falls short of measuring shirking directly, to say nothing of the difficulty of sorting out the direction of causality in these variables. To our knowledge Cappelli and Chauvin (1991), who examine employee dismissal rates using firm-level data, is the only study within the shirking literature to directly measure malfeasant employee behaviour. But even here, the analysis comes up short, since dismissal rates are driven not only by the extent of shirking activity within a workplace, but potentially also by the willingness or ability of employers to detect this behaviour. One can think of many reasons why employers who pay higher wages, for reasons beyond their control, may also have greater incentive to detect or punish shirking activity. But of course, the resulting link between wages and dismissal rates, has little to do with the efficiency wage hypothesis.

3.3 Empirical Identification

The model presented in Chapter 1 provides three main propositions that together not only help us to better understand the empirical absenteeism-weather link that has been examined in the literature, but also provide a strategy for contributing some evidence of the efficiency wage hypothesis. The first proposition tells us that marginal improvements in outdoor weather quality lead to increases in sickness absenteeism of the most healthy employees, which suggests the correlation between weather and sickness absenteeism captures a component in sickness absenteeism that is malfeasant in nature. This proposition demonstrates the reason why using overall sickness absenteeism to proxy shirking is problematic, and justifies our empirical strategy used in this chapter to identify shirking. The second proposition predicts that the marginal effect of weather on sickness absenteeism is larger where existing shirking incentives are low, that is where job acquisition rates are low, sick pay is less generous, and probability of being dismissed when shirking is high. The third proposition, however, predicts that if employers augment the shirking incentives of employees through wage adjustment, that is pay efficiency wages, the effect of the incentive variables mentioned on absenteeism-weather is attenuated, and in the case of sick pay, reversed. Therefore, comparing the results of proposition 2 and 3 provides a strategy for obtaining some evidence of efficiency wages. Specifically, since more generous sick pay is associated with a higher weather-absenteeism relation when efficiency wages are paid, but a lower relation if they are not, an estimate of the sign of this interaction effect offers evidence of the efficiency wage hypothesis. Not only does our focus on the empirical absenteeism-weather relation arguably provide a more compelling measure of shirking activity, but our approach also avoids reliance on trying to estimate the wage premium, which is fraught with complication.

To empirically test the propositions regarding shirking behaviours in Chapter 1, the main strategy used in this Chapter is to link, at the level of the individual employee, data on sickness absenteeism and local weather conditions and examine how this relation varies across employees facing different shirking incentives. The weather information we employ comes from Environment Canada's National Climate Data and Information Archive (NCDIA). Instead of analyzing the separate effects of individual weather elements – temperature, relative humidity, precipitation, wind speed, and cloud cover – this Chapter exploits the weather quality index constructed in Chapter 2, which converts all five weather elements into a single dimension value. The advantage of using the weather quality index is not only that it can measure the joint effect of various weather elements, but also that the weather index itself is constructed so as to capture the substitution in time allocation of workers towards high-utility outdoor recreational activities. Recall also that this weather index is constructed using information on the weekend and weekday evening activities of employees with regular day shifts. We can therefore be sure that the index estimation is not being driven by the shirking decisions of employees.

The data on sickness absenteeism come from Canada's monthly Labour Force Survey (LFS). These data have three important advantages for the purposes of this study. First, although surveys suggest faking sick days is commonplace, the direct causal impact of the weather on overall rates of sickness absenteeism is almost certainly very small.[3] Therefore, large amounts of data are needed in order to identify it with any meaningful precision. In order to limit the possibility that weather fluctuations are directly affecting employee health, we limit our analysis to the period from April to October, which reduces the sample size by nearly half. However, pooling April to October monthly LFS files between 1997 and 2008 provides us with a sample of 1.8 million employees currently employed on a full-time basis in one of the 56 cities for which the local weather data is available. Second, unlike the equivalent survey in the U.S. – the Current Population Survey (CPS) – the LFS identifies not only the usual (contractual) and actual weekly hours worked of employed respondents in the survey reference week (the week containing the 15^{th} day), but also the main reason

[3] For example, a recent online survey by Careerbuilder.com found that one-third of 6,800 employees surveyed had called in sick with a fake excuse at least once over the past year.

for absence in cases where actual hours fall below usual hours. With this information, sickness absenteeism can be distinguished from other types of absenteeism, which may be legitimately influenced by the weather, such as vacations and inclement weather. Lastly, and perhaps most important, the greater the variation in the weather, the more likely it is to provide sufficient utility increments to induce shirking absenteeism. In this sense, Canada offers a more ideal setting to study the absenteeism-weather correlation than more temperate U.S. and European climates.

The main limitation of the LFS data, however, is that only total hours absent in the survey reference week is reported, as opposed to daily or hourly absenteeism. Nonetheless, weather patterns demonstrate substantial serial correlation, that is weather patterns tend to persist over periods longer than a day, so substantial variation exists even when weather conditions are aggregated over a week. In addition, the weather data are observed hourly, so this study is able to examine the differential effect of, for example, good weather on a Friday compared to a Monday. In the baseline case, however, we focus on the average unweighted value of the weather quality index from 9am to 5pm between Monday and Friday is used in the baseline case.

To avoid possible direct effects of the weather on the marginal disutility of work, which would tend to attenuate the estimated absenteeism-weather relation, the sample is further restricted to workers primarily employed indoors using 4-digit industry and occupation codes. The largest groups of workers excluded are those employed in the primary resources, construction, and transportation industries. The Appendix contains a complete list of the groups excluded. Since these codes are only available for each respondent's main job (the job in which they work the most hours), but the absenteeism data identify hours of absence from all jobs, multiple job holders are also excluded from the sample. These restrictions leave us with a final sample of 1,823,074 employees.

Since illegitimate sickness absences are more likely to be short-term, the analysis of the LFS data begins by distinguishing absences that are 8 hours or less in duration from longer part-week or full-week absences. Then, the short-term absence is further distinguished between three reported reasons: (i) own illness or other personal reasons (taking care of kids, elderly people, and other family responsibilities); (ii) vacation; and (iii) other reasons (labour dispute; temporary layoff; holiday; weather; job started or ended during week; working short-time; maternity leave; or other reason). A positive relation between the weather index and the incidence of short-term personal absences (reason (i)) is taken as evidence of shirking absenteeism. Although it is possible to distinguish own illnesses from other personal reasons in the data, attributing an illegitimate absence to a child's illness is arguably no less malfeasant than attributing it to one's own illness. Attributing absence to the illness of family members may in fact be less risky, since it is presumably more difficult to detect. In fact, the results tend to support this conjecture, as absences due to family responsibilities are somewhat more strongly related to the weather than own illness.

A potential concern with the strategy of identifying shirking in this study is that it is capturing implicit agreements between supervisors and employees to use contractual sick days as de-facto vacation days. There is no doubt that these types of agreements exist, but what needs to be true here is that they increase in incidence when outdoor weather conditions improve. In this case, one would clearly have to infer malfeasance on the part of the supervisor, since condoning the use of sick days for the purpose of enjoying good weather is surely not an accepted personnel policy anywhere. It also seems doubtful that workers who use sick days as de-facto vacation days would report the absence to Statistics Canada as resulting from illness. As it turns out, the weather-absenteeism relation identified in the data almost exclusively reflects the behaviour of hourly-paid employees. Since these are exactly the types of employees who are least likely to have contractual sick days, the results also suggest that the weather-sickness absenteeism correlation is not identifying implicit agreements.

To test the propositions in Chapter 1, workers facing different incentives to shirk need to be distinguished. To do this, four covariates are defined to represent the shirking incentives the model in Chapter 1 is concerned about. First, exploiting the rotating sampling structure of the LFS, in which respondents are potentially resampled for 6 consecutive months, the probability of an unemployed worker transitioning to employment in the following month is estimated using a probit model conditional on his or her education, age, duration of current ongoing unemployment spell, month, and city. The unemployment-employment transition probabilities are then predicted at the individual-level and used as to identify the influence of job acquisition rates on the weather-absenteeism relation. Second, to proxy the generosity of sick pay, information on whether respondents are paid on an hourly basis or salaried is employed. Analysis of data from the 1995 Canadian Survey of Work Arrangements shows that, even after conditioning on gender, education, union status, industry, occupation, and geography, hourly paid workers are significantly less likely than salaried workers to be entitled to paid sick leave, which provides some assurance for using this variable as a proxy for sick pay (See Table 3.8 for detailed results). Third, although the theoretical model defines the parameter α as simply a detection probability, its interpretation can be straightforwardly extended to be the joint probability of detection *and* dismissal given detection. To capture variation in the latter probability – the probability of dismissal – two variables are exploited. First, since unionized workers are more likely to have access to a formal grievance process, they are expected to face a lower dismissal probability. Second, typically probationary periods for new employees in Canada are 3 months in duration. Following Ichino and Riphahn (2005), the job tenure data available in the LFS is used to identify a discontinuity in absence behavior at 3 months when job protections usually kick in.

Distinguishing short- and long-term absences, as well as three alternative reasons for short-term absence, a multinomial logit model is employed for the regression analysis.

Specifically, the probability of absence for reported reason j is modeled as:

$$Prob(absence_{ict} = j) = \frac{\exp(\mu_{ictj})}{1 + \exp(\mu_{ict1}) + \cdots + \exp(\mu_{ict4})} \tag{3.3.1}$$

where $j = 1, \ldots, 3$ are personal, vacation, and other reason for short-term absence, respectively; $j = 4$ is a long-term absence; and μ_{ict0} is normalized to zero, so that no absence during the survey reference week ($j = 0$) is the reference category. The linear index μ_{ictj} for $j = 1, \ldots, 4$ are specified as follows:

$$\begin{aligned}\mu_{ictj} = \Big[& f_j(weather_{ct}) + \theta_{1j} ar_{ict} + \theta_{2j} hr_{ict} + \theta_{3j} un_{ict} + \theta_{4j} ten_{ict} + \theta_{5j} ten^2_{ict} + \theta_{6j} \mathbf{1}[ten_{ict} \geq 3] + \\ & \mathbf{dem}_{ijt}\lambda_{1j} + \mathbf{ind}_{ict}\lambda_{2j} + \mathbf{occ}_{ict}\lambda_{3j} + \mathbf{z}_c\lambda_{4j} + \mathbf{x}_t\lambda_{5j} \Big].\end{aligned} \tag{3.3.2}$$

where $weather_{ct}$ is the average value of the weather quality index between 9am and 5pm from Monday to Friday in city c in week t; ar_{ict} is the job acquisition rate; hr_{ict} and un_{ict} are dummies indicating hourly-paid and unionized, respectively; ten_{ict} is months of job tenure; $\mathbf{1}[ten_{ict} \geq 3]$ is an indicator function identifying a discontinuity in absence probabilities at 3 months of job tenure; $\mathbf{dem}_{ijt}$ is a vector of individual demographic variables consisting of a quartic form in age, eight education categories and gender. $\mathbf{ind}_{ict}$, $\mathbf{occ}_{ict}$, $\mathbf{z}_c$, and $\mathbf{x}_t$ are vectors of industry, occupation, city, and month dummies, respectively. The only remaining issue is how to specify the weather function $f_j(\cdot)$. The theory implies that the marginal effect of the weather on the probability of absence is nonlinear. Estimating the function using various polynomials, the results show that the functional form is clearly nonlinear. However, as is discussed in detail in the following section, little explanatory power is gained beyond a function with quadratic form.[4]

Propositions 2 and 3 in Chapter 1 are concerned with the marginal effect of the weather across other shirking incentive parameters. To test these propositions, equation (3.3.2) is also estimated including interactions of the weather function $f_j(\cdot)$ with either the job acquisition rate (ar_{ct}), the hourly-rate dummy (hr_{ict}), the union dummy (un_{ict}), and the post-probation dummy ($\mathbf{1}[ten_{ict} \geq 3]$). In the absence of efficiency wages (Proposition 2), it is expected that the weather effect is highest where shirking incentives are lowest implying that the ar_{ct} interaction is negative; the hr_{ict} interaction is positive; the un_{ict} interaction is negative; and the $\mathbf{1}[ten_{ict} \geq 3]$ interaction is negative. However, to the extent that employers use efficiency wages to augment shirking incentives, the hr_{ict} interaction should be negative, as employers use wages to limit the relative shirking incentives of salaried workers.

[4]The main difference in adding higher-order polynomials is the estimated marginal effect of the weather on short-term personal absences becomes very flat, rather than declining, at the upper tail of the weather quality distribution.

While the model provides strong justification for using the empirical weather-absenteeism relation to identify one particular type of shirking activity, the covariates used to proxy shirking incentives, most notably the hourly-paid indicators, might be to some extent endogenous. In particular, we are unable to rule out the possibility that rather than using wages to influence shirking incentives, employers may in fact use payment arrangement, that is hourly versus salaried, which is assumed exogenous in the theoretical model, to influence these incentives (MacLeod and Malcomson, 1998). Unfortunately, no credible instrumental variables for the hourly-paid indicator are available. We instead perform two robustness checks using additional controls intended to capture the underlying heterogeneity that could bias our estimates. First, a wage premium is estimated for each individual in the data by regressing their effective hourly wage provided in the LFS data on eight education categories, a quartic in age, and a vector of city dummies. Then equation (3.3.2) is estimated including and excluding the residual from this wage regression as an additional regressor to examine whether the estimated interaction effects change significantly, implying that wages are being used by employers to augment shirking incentives. Second, exploiting the rotating sampling structure of the LFS, in which respondents are potentially resampled for 6 consecutive months, we are able to identify past absenteeism for personal reasons at the individual level. Specifically, we restrict our sample to individuals observed over the entire 6-month panel, and construct a lagged absenteeism variable in period t equal to the percentage of the previous 5 months in which the individual reported a short-term absence due to personal reasons. By restricting our sample to only those observations that are in the sixth month of their rotation, we severely reduce the our overall sample size. However, our hope is that including this lagged dependent variable as an additional regressor in estimating equation (3.3.2) will address any bias resulting from unobservable health differences between workers that are correlated with our shirking incentive proxies.

3.4 Results

We begin our analysis of the LFS data by comparing unconditional sample mean probabilities of absence across the key covariates thought to influence shirking incentives. The results are presented in Table 3.1. Comparing the incidence of a short-term absence for personal reasons across quintiles of the sample weather quality distribution, there is a clear tendency for personal absenteeism to rise with good weather. Although as expected the magnitude of the effect appears very small, as a percentage change, it is a 33% increase in probability from 0.0220 to 0.0295. The unconditional variation in the weather is, however, overwhelmingly seasonal. That is, most of the variation in the weather data is between days in April and October, when temperatures tend to be lower, and in July and August, when they are on average higher. This seasonality in the weather almost certainly accounts for much of the tendency for short-term vacation absences to rise with weather quality.

This seasonality in the weather likely also accounts for the strong decline in short-term absences for other reasons, as this category includes absences due to holidays, as Easter and Thanksgiving both potentially fall in the survey references week in April and October, respectively, when weather quality is on average poor. The seasonality in weather can also explain the tendency for long-term absences to increase with the quality of the weather because vacations in excess of one day are most likely in July and August when the weather is best. Since genuine health status may similarly vary with season, it is important that the weather-absenteeism relation we identify is conditional on the calendar month.[5]

The differences in short-term personal absences across job acquisition rates, union status, probation status, and estimated wage premiums are all consistent with the expected shirking incentive effects of these variables – personal absenteeism is higher when job acquisition rates are high, when job protection is high, and when wage premiums are low. A higher incidence of personal absence among hourly-rate workers is, however, unexpected, given that they are less likely to be paid for time off. Of course, it is unclear to what extent any of these differences reflect genuine health. For example, a wage premium may induce less reported sickness absenteeism, but it could also be a consequence of good health. And workers paid an hourly-rate may be on average less healthy for reasons that have nothing to do with shirking incentives. It is precisely the difficulty in interpreting these differences as shirking incentives that motivates our use of the weather-absenteeism relation.

Table 3.2 presents the results from estimating the baseline multinomial logit model defined by equations (3.3.1) and (3.3.2). The main finding is that the weather quality index only appears to have a statistically significant effect on the incidence of short-term personal absences. However, this result is somewhat sensitive to the choice of function form. When we instead estimate a linear weather effect, the weather also has a statistically significant effect on vacation absences. Figure 3.1, shows changes in the predicted probability of short-term absences due to personal reasons from the 5th to 95th percentile of the weather quality distribution separately using a linear, quadratic, cubic and quartic weather function. The results show that there is clearly a nonlinear effect of weather, but adding a cubic or quartic term does little to change the estimated effect. We, therefore, report results from using the quadratic function. The results in Table 3.2 show that the direction of the weather effect on personal absences is consistent with Proposition 1, that is weather conditions more conducive to outdoor recreation result in more sickness absenteeism. The marginal effect is, decreasing, but positive up to an index value of 0.162, which falls above the 90$^{\text{th}}$ percentile of the weather quality distribution. At the mean of the data, a one standard deviation increase in the weather quality index increases the probability of a short-term absence for personal reasons from 0.0299 to 0.0318. This is clearly not a large effect, but given the substantial sample size, it is statistically significant at the 5% level.

[5]Note that there is no obvious pattern in the rates of short-term sickness absences across months in Table 3.1, suggesting that at least over these non-winter months, there is little seasonality in health.

Table 3.3 reports the results from estimating the same model as in Table 3.2, but distinguishing short-term absences due to one's own sickness and due to the sickness of other family members. As speculated, the effect of weather on the absence due to taking care of sick family members is even stronger. Given that using the excuse of taking care sick family members to avoid work duty is less likely to be detectable than using own sickness, the stronger correlation suggests that it is indeed the shirking activity we are identifying here.[6]

Before considering how this weather effect on personal absences varies across employees facing different shirking incentives, the effect of weather was further explored in two ways. First, rather than regressing on the overall average daytime weather quality between Monday and Friday, the marginal effect of daytime weather is examined separately for each day, conditional on the daytime weather of the other four working days of the week. The results are presented in Table 3.4. Although most of the individual weather terms appear to be insignificant, joint tests of the significance of the linear and quadratic weather terms are performed separately. The results suggest that personal absenteeism is most sensitive to the weather on Tuesday. It may be, at first, surprising that Tuesday weather is more influential than the weather on Fridays or Mondays, given that employees are more likely to be tired on Fridays as well as more likely to use either the Friday before or the Monday after a weekend for a sufficient block of time to enable certain activities, such as traveling. However, perhaps as a result of these reasons, employee absenteeism on Mondays and Fridays is less likely to be credible, and employees are therefore less likely to misreport sickness on these days. The result of Tuesday's weather being most influential to absenteeism is also found when only a linear functional form of the weather index variable is used in the analysis. Interestingly though, Monday weather appears to influence other types of short-term absences, as well as long-term absences, although except for vacations, the estimated marginal effects are actually decreasing over most of the weather distribution (the negative quadratic term tends to dominate). This suggests that the results for Mondays are capturing higher absenteeism on days other than Monday during weeks when Monday's weather is poor relative to the average weather during the rest of the week.

Following on this idea that the marginal utility of the weather is higher when recent past weather has been worse, Table 3.5 presents the results from interacting the quadratic weather function with the average daytime weather index on the weekend preceding the survey reference week. As in Table 3.2, the current weather appears only to affect personal absenteeism and not other types of absenteeism. Moreover, the interaction of current and past weather suggests that past weather influences the marginal utility of the current

[6] All the regression analysis in this study has been conducted with both considering one's own sickness and family members' sickness as same type of absence and as different types of absence. All the results appear to be that the effect of weather is stronger on short-term absence due to family members' sickness than due to one's own sickness.

weather. Specifically, the estimates suggest that conditional on average weekend weather of 0.05 (roughly the 15$^{\text{th}}$ percentile), the probability of a personal absenteeism increases by 5.84% (from 0.0262 to 0.0278) when the average workweek weather increases from 0.05 to 0.06. When average weekend weather is 0.15 (roughly the 90$^{\text{th}}$ percentile), on the other hand, a one-point increase in average workweek weather (from 0.15 to 0.16) increases personal absenteeism by only 2.01% (from 0.0317 to 0.0324).[7]These results provide evidence of intertemporal substitution of leisure, corroborating the findings of Connolly (2008). The difference here is that Connolly's data does not allow her to distinguish the types of absenteeism, whereas our results imply the intertermporal substitution is at least partially malfeasant in nature.

Finally, Table 3.6 presents the results from interacting the quadratic weather function with the job acquisition rate; hourly-paid indicator; union indicator; and post-probation indicator. Specifications (1) and (2) present the results from excluding and including, respectively, the estimated wage premium as an additional regressor. Due to the quadratic specification, the ranking of the magnitude of the marginal effects potentially changes over the empirical support of the weather quality index. To make the results more transparent, in Figure 3.2 we plot the predicted probabilities of personal absence (at the mean of the data) between the 5$^{\text{th}}$ and 95$^{\text{th}}$ percentiles of the weather distribution separately for each of the four interacting variables. In the case of the job acquisition rate, this is done by comparing the profile between a job acquisition rate at the 25$^{\text{th}}$ and 75$^{\text{th}}$ percentiles (0.29 compared to 0.39). Last, in cases where the ranking of the marginal effects switches, the vertical lines indicate the point at which the marginal effects are equal in magnitude.

In the case of both the job acquisition rate and unionization status, the estimated interactions are statistically insignificant in the linear term but significant in the quadratic term, although the point estimates for the unionization dummy are the right sign over most of the weather quality distribution. Specifically, the point estimates imply that the marginal effect of the weather is bigger for non-unionized workers beyond the 30$^{\text{th}}$ percentile of the weather quality distribution (an index value of roughly 0.071). Based on Proposition 3, if the efficiency wages are paid, these interaction effects will tend to be attenuated, so these results are not inconsistent with the efficiency wage hypothesis. However, as a robustness check, conditioning on the wage premium, does essentially nothing to change these results. Therefore, although there does exist a clear positive relation between overall personal absenteeism and job acquisition rates, the results do not support the idea that, in choosing to shirk contractual work hours, workers think about the likelihood of finding a replacement job in the event that their malfeasance is detected and they are dismissed.

[7]The effect of weather of the past weekend is also estimated with a linear weather function. In this case the linear weather variable is positive and significant and the interaction variable is negative and marginally significant, implying the marginal utility of workweek weather is increasing with the adverseness of the previous weekend weather over the entire distribution of the workweek weather.

Rather the results are more consistent with the procyclical relation between health and the business cycle, which has been documented elsewhere, reflecting genuine health variation, as opposed to shirking incentives.

A cleaner test of the efficiency wage hypothesis is with regard to the effect of sick pay on shirking incentives. Using hourly-paid status to proxy sick pay, the estimates point to a higher marginal effect of the weather for hourly-paid workers above the 35^{th} percentile (an index value of about 0.078) of the weather quality distribution. Below this point the marginal effects are virtually identical for hourly-paid and salaried workers. Going from the median weather quality (about 0.095) to the 95^{th} percentile, has essentially no effect on the short-term personal absenteeism rate of salaried employees, while it increases the rate for hourly-rate employees from about 0.038 to 0.044, a 17% increase. This result is consistent with Proposition 2 – that hourly workers who are receiving less generous sick pay face a lower existing shirking incentive and will be more responsive to weather improvements. However, this result is not consistent with employers using wages to influence shirking incentives, since according to Proposition 3, we should then see a greater responsiveness to weather improvements among workers with more sick pay(salaried workers). Moreover, the results are virtually identical whether or not we condition on the wage, suggesting further that wage rates are not influencing shirking incentives.

Lastly, the results in the final two columns of Table 3.6, point to very different responses to weather improvements between pre- and post-probation employees, but the implications for efficiency wage theory is more mixed. Specifically, marginal weather improvements below the median weather quality (an index value of 0.095) have a bigger impact on the personal absence rates of post-probation employees, whereas weather improvements above the median weather quality have a larger impact on pre-probation employees. Assuming a higher probability of dismissal among pre-probation employees, the results below the median are therefore consistent with Proposition 3, and the payment of efficiency wages, while the results above the median are consistent with both Proposition 2 and 3. One could argue that the types of outdoor activities that employees are imagined to substitute towards when they skip work, are more likely to be induced by marginal weather improvements at the upper end of the distribution, in which case the result is not particularly informative. We might expect, for example, that going from bad to not so bad weather induces less workday activity on the golf course, than going from good to great weather. However, once again, virtually identical results are obtained whether or not the analysis is conditioned on the wage premium, suggesting once again that wages are not influencing the relative shirking incentives of pre- and post-probationary employees.

Table 3.7 presents the results of the robustness analysis when employees' lagged short-term absence due to personal reasons is included as a control variable to overcome the potential bias caused by the correlation between the proxy variables for shirking incentives and unobserved individual heterogeneity in health. Due to the substantial decrease in sample size, the statistical significance of the estimates of the interaction effects drops substantially, but the point estimates in general remain similar. As in the results using the full sample, controlling for the wage premium does essentially nothing to change the estimates, which, again, is inconsistent with the efficiency wage hypothesis. In terms of the changes in the responses of employees facing different shirking incentives over the empirical support of the weather quality index, the results are also comparable with what is shown in Figure 3.2. In comparing hourly-paid and salaried employees, the marginal effect of weather on hourly-paid employees becomes bigger when weather index is greater than 0.069. In comparison to the turning point before, which is 0.078, this suggests that over an even bigger range of weather quality, hourly paid employees are more responsive to weather quality improvement taking sick absence. The effect of weather on non-unionized employees and pre-probation workers becomes bigger when weather quality is greater than 0.099 and 0.111, respectively. Although the turning point, compared with before, appears at larger values of the weather index, the overall profiles of the weather effects on employees facing different shirking incentives have not changed. Therefore, the robustness analysis again, provides evidence supportive of Proposition 2, but not of Proposition 3.

3.5 Summary

By empirically testing the three main propositions in Chapter 1, this chapter furthers our understanding of employee shirking behaviour, especially its response to shirking incentives. 12 years of employee data from Canada's monthly Labour Force Survey (LFS), which contains information on incidences and reasons of employee absence, are linked with weather quality index, which gives us a sample of 1.8 million indoor employees. We then examine the relation between short-term absence due to sickness reasons and weather quality between April and October and how this relation varies across workers facing different incentives.

We find clear evidence of a positive relation between weather and sickness absenteeism, and according to Proposition 1 in Chapter 1, this correlation identifies a shirking component in the overall reported sickness absenteeism. With this more precisely measured shirking activity, we examine how employees shirking behaviour responses to shirking incentives, in particular the generosity of sick pay, the probability of replacing a lost job, and the probability of dismissal if getting caught shirking. Four variables – hourly paid, predicted job acquisition rate, and union status and tenure – are used to proxy these three shirking incentives, respectively. We find employees' responses to weather improvements are larger when existing shirking incentives are low, which is among hourly paid, non-unionized, and employees in the probation period. Our results in general support the Proposition 2 in Chapter 1, and appears inconsistent with the efficiency wage hypothesis.

Table 3.1: Sample mean probabilities of absence

	Short-term absence			Long-term absence
	Personal	Vacation	Other	
Weather				
Fifth quintile	0.0295	0.0244	0.0053	0.1656
Fourth quintile	0.0312	0.0218	0.0089	0.1581
Third quintile	0.0312	0.0223	0.0210	0.1299
Second quintile	0.0297	0.0185	0.1042	0.1097
First quintile	0.0220	0.0174	0.2606	0.1250
Job acquisition rate				
Above median	0.0305	0.0227	0.0603	0.1456
Below median	0.0275	0.0193	0.1017	0.1275
Pay status				
Salaried	0.0270	0.0271	0.1021	0.1482
Hourly paid	0.0300	0.0159	0.0717	0.1251
Union status				
Unionized	0.0326	0.0228	0.0894	0.1885
Non-unionized	0.0272	0.0198	0.0828	0.1142
Probation status				
Over 3 months	0.0288	0.0216	0.0861	0.1415
Under 3 months	0.0279	0.0089	0.0664	0.0544
Wage premium				
Above median	0.0284	0.0248	0.0936	0.1419
Below median	0.0290	0.0165	0.0757	0.1280
Month				
April	0.0296	0.0186	0.1478	0.1078
May	0.0337	0.0327	0.0086	0.0937
June	0.0374	0.0250	0.0053	0.0971
July	0.0260	0.0294	0.0041	0.2191
August	0.0254	0.0295	0.0129	0.2271
September	0.0361	0.0187	0.0056	0.0974
October	0.0114	0.0198	0.4685	0.1490

Notes: Short-term absence is defined as total hours absent of 8 hours or less. Weather is the average value of the weather quality index from Monday to Friday between 9am and 5pm.
Source: 1997 to 2008 Labour Force Survey (LFS) and Canadian National Climate Data and Information Archive (NCDIA).

Table 3.2: Multinomial logit estimates of the probability of absence from work during survey reference week

	Short-term absence			Long-term absence
	Personal	Vacation	Other	
Weather	5.8397**	2.4065	-6.6625	-0.7192
	(2.5311)	(3.6250)	(14.0593)	(3.0827)
Weather^2	-18.0552**	-1.3967	-7.3017	7.9204
	(8.7856)	(13.8529)	(60.9238)	(11.5980)
Job acquisition rate	5.2307***	6.4584***	5.6673**	2.6802***
	(0.5097)	(0.6354)	(2.4843)	(0.5120)
Hourly paid	0.1082***	-0.1827***	-0.1687***	-0.0679***
	(0.0154)	(0.0184)	(0.0160)	(0.0092)
Unionized	0.2261***	0.1134***	0.0809***	0.2931***
	(0.0156)	(0.0181)	(0.0150)	(0.0087)
Tenure	0.0004*	0.0043***	0.0017***	0.0041***
	(0.0002)	(0.0002)	(0.0002)	(0.0001)
Tenure^2/100	-0.0001	-0.0001	-0.0003***	-0.0007***
	(0.0001)	(0.0001)	(0.00004)	(0.00003)
Tenure over 3 months	0.0997***	0.5259***	0.1320***	0.7313***
	(0.0256)	(0.0391)	(0.0218)	(0.0194)
Pseudo R^2			0.1700	
N			1,822,726	

Notes: Short-term absence is defined as total hours absent of 8 hours or less. Weather is the average value of the weather quality index from Monday to Friday between 9am and 5pm. Standard errors are clustered by city and month. Regression also controls for age, gender, education, city, month, industry and occupation. ***,** ,* indicate statistical significance at the 1%, 5%, and 10% levels, respectively.
Source: 1997 to 2008 Labour Force Survey (LFS) and Canadian National Climate Data and Information Archive (NCDIA).

Table 3.3: Multinomial logit estimates of the probability of absence from work during survey reference week (sperate short-term absence due to taking care of sick family members)

	Short-term absence				Long-term absence
	Own sickness	Family responsibility	Vacation	Other	
Weather	3.6105	9.4769***	2.4071	-6.6657	-0.7219
	(2.6429)	(3.1654)	(3.6250)	(14.0599)	(3.0828)
Weather^2	-11.4345	-28.4845***	-1.3981	-7.2992	7.9275
	(9.3007)	(11.0172)	(13.8530)	(60.9277)	(11.5986)
Job acquisition rate	6.6096***	3.1412***	6.4517***	5.6722**	2.6829**
	(0.5349)	(0.6183)	(0.6350)	(2.4842)	(0.5120)
Hourly paid	0.0816***	0.1526***	-0.1827***	-0.1688***	-0.0679***
	(0.0193)	(0.0226)	(0.0184)	(0.0160)	(0.0092)
Unionized	0.3354***	0.0344	0.1133***	0.0812***	0.2932***
	(0.0183)	(0.0257)	(0.0181)	(0.0150)	(0.0087)
Tenure	0.0003	0.0006*	0.0043***	0.0017***	0.0041***
	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0001)
Tenure^2/100	-0.0005	-0.0001	-0.0007***	-0.0003***	-0.0007***
	(0.0001)	(0.0001)	(0.0001)	(0.00004)	(0.00003)
Tenure over 3 months	0.1462***	0.0187	0.5259***	0.1322***	0.7313***
	(0.0301)	(0.0468)	(0.0391)	(0.0218)	(0.0194)
Pseudo R^2			0.1669		
N			1,822,726		

Notes: Short-term absence is defined as total hours absent of 8 hours or less. Weather is the average value of the weather quality index from Monday to Friday between 9am and 5pm. Standard errors are clustered by city and month. Regression also controls for age, gender, education, city, month, industry and occupation. ***,** ,* indicate statistical significance at the 1%, 5%, and 10% levels, respectively.
Source: 1997 to 2008 Labour Force Survey (LFS) and Canadian National Climate Data and Information Archive (NCDIA).

Table 3.4: Multinomial logit estimates of the probability of absence from work during survey reference week by daily weather quality

	Short-term absence			Long-term absence
	Personal	Vacation	Other	
Monday-weather	1.1701	11.5961***	41.6121***	9.7619***
	(2.0197)	(3.0956)	(11.5863)	(2.6812)
Monday-weather^2	-5.6987	-44.5537***	-207.4685***	-37.0173***
	(7.8533)	(12.2276)	(62.2020)	(10.5830)
Tuesday-weather	0.8708	-0.4580	18.2635	1.1884
	(2.4264)	(3.3008)	(11.8101)	(2.5118)
Tuesday-weather^2	4.7046	7.8856	-96.1747	-0.3330
	(9.1755)	(12.6568)	(61.4790)	(9.9841)
Wednesday-weather	-2.3647	-5.8366**	-14.8015	-2.7336
	(2.2124)	(2.7882)	(13.3083)	(2.4308)
Wednesday-weather^2	6.8555	20.6834*	61.9287	8.8665
	(8.6711)	(10.9077)	(66.6439)	(9.5945)
Thursday-weather	3.5731*	-4.3381	-20.3523	-4.4294
	(2.1574)	(3.1314)	(14.7574)	(2.9146)
Thursday-weather^2	-13.1702	19.1868	111.2558	20.6195*
	(8.5848)	(12.2611)	(75.0729)	(11.5245)
Friday-weather	2.2493	2.6537	-13.7873	-3.0255
	(2.2325)	(2.6220)	(11.0951)	(2.2952)
Friday-weather^2	-9.9515	-8.4189	36.4929	10.8657
	(8.7052)	(10.1489)	(52.8644)	(9.0207)
Pseudo R^2			0.1730	
N			1,821,554	

Notes: Short-term absence is defined as total hours absent of 8 hours or less. Weather is the daily average value of the weather quality from 9am and 5pm. Standard errors are clustered by city and month. Regression also controls for age, gender, education, city, month, industry and occupation. ***,** ,* indicate statistical significance at the 1%, 5%, and 10% levels, respectively.
Source: 1997 to 2008 Labour Force Survey (LFS) and Canadian National Climate Data and Information Archive (NCDIA).

Table 3.5: Multinomial logit estimates of the probability of absence from work during survey reference week conditional on previous weekend weather

	Short-term absence			Long-term absence
	Personal	Vacation	Other	
Weather	13.4784***	4.6701	-4.3531	-2.8219
	(4.3604)	(6.5832)	(25.1010)	(5.8682)
Weather^2	-65.0444***	-19.6809	-24.7849	22.8660
	(23.2283)	(34.4317)	(150.5047)	(30.4340)
Weather*Previous weekend weather	-50.3216**	0.0845	-20.1847	5.6590
	(21.8373)	(25.9500)	(104.1999)	(21.5950)
Weather^2*Previous weekend weather	355.3282**	72.6816	156.3359	-76.3461
	(143.6966)	(179.1534)	(837.6791)	(153.1467)
Pseudo R^2			0.1701	
N			1,810,8911	

Notes: Short-term absence is defined as total hours absent of 8 hours or less. Weather is the average value of the weather quality index from Monday to Friday between 9am and 5pm. Standard errors are clustered by city and month. Regression also controls for age, gender, education, city, month, industry and occupation. ***,**,* indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Source: 1997 to 2008 Labour Force Survey (LFS) and Canadian National Climate Data and Information Archive (NCDIA).

Table 3.6: Multinomial logit estimates of the probability of short-term absence due to personal reasons

	Interaction variable							
	Job acquisition rate		Hourly paid		Unionized		> 3 months tenure	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Weather	15.3496**	15.3186**	8.9173***	8.9041***	5.4015**	5.3927**	-0.1290	-0.1379
	(6.3187)	(6.3194)	(2.8543)	(2.8543)	(2.5703)	(2.5702)	(3.2427)	(3.2429)
Weather^2	-74.3030***	-74.1149***	-37.0017***	-36.9083***	-14.7271	-14.6819	12.6539	12.6847
	(27.6820)	(27.6830)	(10.7606)	(10.7601)	(9.0467)	(9.0467)	(12.4938)	(12.4949)
Weather*Interaction	-23.7817	-23.7099	-4.3134***	-4.2984***	1.6340	1.6559	6.5420***	6.5483***
	(14.4926)	(14.4910)	(1.5015)	(1.5002)	(1.4381)	(1.4387)	(2.4676)	(2.4679)
Weather^2*Interaction	140.4293**	139.9816**	27.3392***	27.2084***	-11.3591*	-11.4818*	-33.6835***	-33.7024***
	(64.4837)	(64.4727)	(7.2115)	(7.2056)	(6.6879)	(6.6916)	(10.7260)	(10.7281)
Job acquisition rate	6.0343***	6.0182***	5.2331***	5.2195***	5.2281***	5.2138***	5.2307***	5.2165***
	(0.8941)	(0.8942)	(0.5100)	(0.5097)	(0.5095)	(0.5092)	(0.5095)	(0.5093)
Hourly paid	0.1084***	0.1146***	0.2420***	0.2479***	0.1085***	0.1148***	0.1083***	0.1146***
	(0.0154)	(0.0155)	(0.0724)	(0.0724)	(0.0154)	(0.0155)	(0.0154)	(0.0155)
Unionized	0.2262***	0.2214***	0.2252***	0.2205***	0.1864***	0.1807**	0.2261***	0.2212***
	(0.0156)	(0.0155)	(0.0156)	(0.0155)	(0.0716)	(0.0717)	(0.0156)	(0.0155)
Tenure	0.0004*	0.0004**	0.0004**	0.0004**	0.0004**	0.0004**	0.0004*	0.0004**
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Tenure^2/100	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Tenure over 3 months	0.1000***	0.0993***	0.1004***	0.0998***	0.1000***	0.0993***	-0.1765	-0.1775
	(0.0256)	(0.0256)	(0.0256)	(0.0256)	(0.0256)	(0.0256)	(0.1379)	(0.1379)
Wage premium	–	0.0385**	–	0.0373***	–	0.0392**	–	0.0388**
		(0.0181)		(0.0181)		(0.0181)		(0.0181)
Pseudo R^2	0.1702	0.1707	0.1703	0.1708	0.1703	0.1709	0.1701	0.1707
N	1,822,726	1,822,726	1,822,726	1,822,726	1,822,726	1,822,726	1,822,726	1,822,726

Notes: Short-term absence is defined as total hours absent of 8 hours or less. Weather is the average value of the weather quality index from Monday to Friday between 9am and 5pm. Standard errors are clustered by city and month. Regression also controls for age, gender, education, city, month, industry and occupation. ***,** ,* indicate statistical significance at the 1%, 5%, and 10% levels, respectively.
Source: 1997 to 2008 Labour Force Survey (LFS) and Canadian National Climate Data and Information Archive (NCDIA).

Table 3.7: Multinomial logit estimates of the probability of short-term absence due to personal reasons (conditional on lagged short-term personal absence)

	Interaction variable							
	Job acquisition rate		Hourly paid		Unionized		> 3 months tenure	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Weather	13.6381	16.0039	9.9110*	9.2127*	5.8387	5.6017	-23.6105*	-25.4208**
	(13.2247)	(12.9900)	(5.5605)	(5.3562)	(4.7799)	(4.5700)	(12.6164)	(12.9480)
Weather^2	-68.1225	-77.5830	-36.2868*	-33.7233	-11.7387	-12.2873	113.4201**	119.5946**
	(60.7423)	(59.6836)	(22.3140)	(21.7130)	(17.9585)	(17.2973)	(55.8510)	(57.9497)
Weather*Interaction	-23.4912	-21.5110	-3.0775	-2.6744	5.7712	5.1806	32.0817***	33.4586***
	(32.2682)	(32.2608)	(4.1224)	(4.1383)	(4.3722)	(4.3872)	(12.1564)	(12.5359)
Weather^2*Interaction	112.4648	141.6881	22.2659	19.1225	-28.9107	-25.5496	-138.0124**	-143.6031**
	(149.8534)	(149.4697)	(19.4145)	(19.5146)	(19.8848)	(19.9624)	(54.6416)	(56.7986)
Job acquisition rate	5.9567***	5.6104***	5.8681***	5.0532***	5.8655***	5.0534***	5.8643***	5.0445***
	(1.8364)	(1.8409)	(0.9913)	(0.9709)	(0.9909)	(0.9705)	(0.9900)	(0.9696)
Hourly paid	0.1124**	0.0945**	0.1797	0.1553	0.1122**	0.0942**	0.1129**	0.0950**
	(0.0439)	(0.0445)	(0.2065)	(0.2065)	(0.0438)	(0.0444)	(0.0439)	(0.0444)
Unionized	0.2380***	0.2086***	0.2369***	0.2075***	-0.0167	-0.0241	0.2380***	0.2087***
	(0.0441)	(0.0447)	(0.0441)	(0.0446)	(0.2211)	(0.2221)	(0.0441)	(0.0446)
Tenure	-0.0004	-0.0003	-0.0004	-0.0003	-0.0004	-0.0003	-0.0004	-0.0003
	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0006)
Tenure^2/100	0.0001	0.0001	0.0002	0.0001	0.0001	0.0001	0.0001	0.0001
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Tenure over 3 months	0.0547	0.0425	0.0554	0.0435	0.0543	0.0422	-1.5609**	-1.6461***
	(0.1408)	(0.1420)	(0.1407)	(0.1419)	(0.1406)	(0.1418)	(0.6197)	(0.6344)
Lagged short-term personal absence	–	4.0902***	–	4.0893***	–	4.0900***	–	4.0921***
		(0.1349)		(0.1348)		(0.1348)		(0.1348)
Pseudo R^2	0.1754	0.1801	0.1753	0.1801	0.1754	0.1802	0.1751	0.1799
N	188,797	188,797	188,797	188,797	188,797	188,797	188,797	188,797

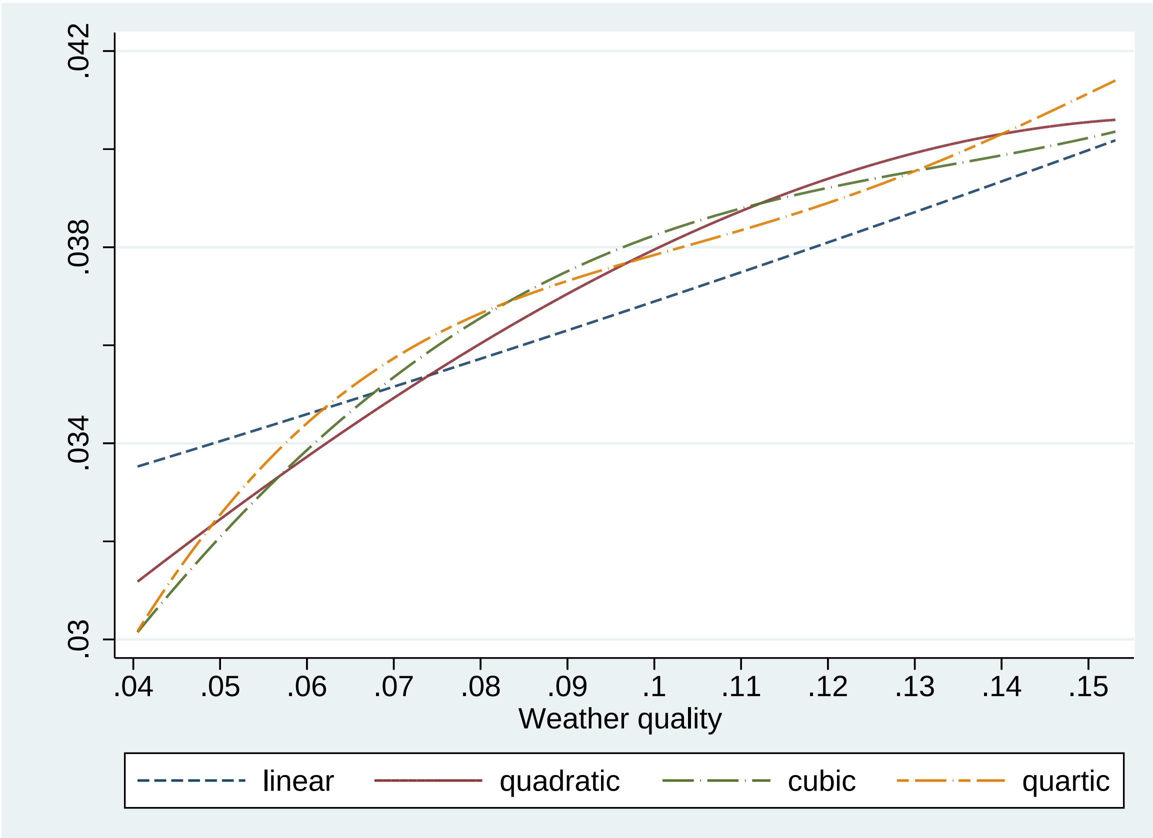
Notes: Short-term absence is defined as total hours absent of 8 hours or less. Weather is the average value of the weather quality index from Monday to Friday between 9am and 5pm. Standard errors are clustered by city and month. Regression also controls for age, gender, education, city, month, industry and occupation. ***,** ,* indicate statistical significance at the 1%, 5%, and 10% levels, respectively.
Source: 1997 to 2008 Labour Force Survey (LFS) and Canadian National Climate Data and Information Archive (NCDIA).

Table 3.8: Probit estimates of correlation between hourly paid employees and the entitlement of sick pay

	Coefficient	Standard error
Hourly	-0.6294***	0.0322
Union	0.2976***	0.0763
Male	0.2448***	0.0324
Age		
17-19	0.2109	0.2614
20-24	0.8803***	0.2520
25-34	1.4250***	0.2500
35-44	1.5350***	0.2499
45-54	1.6094***	0.2504
55-64	1.6001***	0.2540
65-69	0.9423***	0.3172
Education		
Some secondary education	-0.1034**	0.0421
High school	-0.0281	0.0368
Some pose-secondary education	-0.0724	0.0465
University	0.0933**	0.0459
Pseudo R^2	0.2356	
N	20,813	

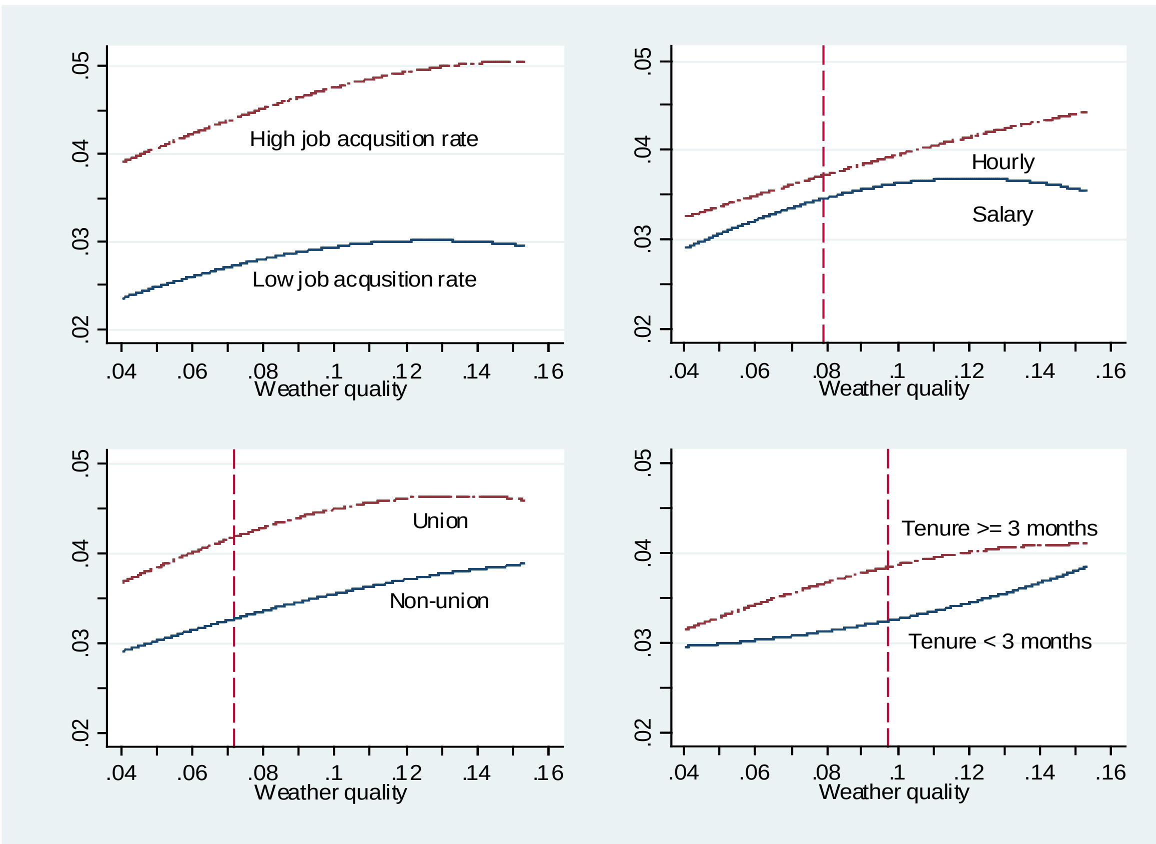
Notes: Regression analysis is conditional on occupation, industry, and province. ***,** ,* indicate statistical significance at the 1%, 5%, and 10% levels, respectively.
Source: 1995 Survey of Work Arrangement.

Figure 3.1: Predicted probability of short-term absence due to personal reasons relative to no absence (odds ratios), by the functional form of weather



Notes: Predicted probabilities are from estimates when weather takes the function form of linear, quadratic, cubic and quartic. In each case, the probabilities are predicted at the means of the sample.

Figure 3.2: Predicted probability of short-term absence due to personal reasons relative to no absence (odds ratios), by shirking incetives



Notes: Predicted probabilities are from estimates in Specification (1) of Table 3.6. In each case, the probabilities are predicted at the means of the sample. High and low job acquisition rates are 0.29 and 0.39, respectively, which are the 25^{th} and 75^{th} percentiles in the sample. The vertical lines indicate the value of the weather index where the slopes of the profiles are equal.

Chapter 4

Why Income Matters for Participation in Leisure Time Physical Activities: Evidence from Weather Shocks

4.1 Introduction

Despite the well-documented benefits of physical activity to physical and mental health, close to half (48%) of Canadian adults are still physically inactive in their leisure time, meaning that they typically do less than the equivalent of one-half hour of walking per day (Gilmour, 2007). The economic burden of this level of physical inactivity is estimated to be substantial. In Canada, the health care costs related to physical inactivity was estimated at \$5.3 billion, 2.6% of total health care costs in 2001 (Katzmarzyk and Janssen, 2004). Sari (2009) estimates that physically inactive individuals, compared with active individuals use 2.4% to 9.6% more health care services. Promoting leisure time physical activities (LTPA) is, therefore, considered a public health priority since physical inactivity is a modifiable risk factor for a number of health problems including cardiovascular disease, osteoporosis, high blood pressure, obesity, diabetes, depression, stress and anxiety (Warburton, Nicol and Bredin, 2006). There, however, exits considerable variations in physical activity participation across different segments of the population. Most notably, research consistently shows that people in lower socioeconomic groups are less likely to engage in physical activities (Ball et al., 2006; Brown and Roberts, 2011; and Droomers, Schrijvers, and Mackenbach, 2001). Because interventions aimed at promoting participation in physical activities are more effective if they address the needs of a particular target group (Booth et al., 1997), understanding the mechanisms through which socioeconomic status affects physical activities is of great importance.[1]

While the empirical correlation between socioeconomic status and participation in physical activities is well established, the underlying mechanisms are not well understood. The extant literature has mainly focused on individual characteristics to explain individuals' decisions regarding physical activities. Similar to difficulties in explaining the socioeconomic gradient observed in many health behaviour, such as smoking, drinking, and dieting, individuals' socioeconomic status in this context, is highly endogenous. Heterogeneity across individuals, much of which is unobservable in the data, makes it difficult to interpret the relation between socioeconomic characteristics and health behaviours. Although several mediating variables, such as attitude, self-efficacy, and self-schemata, are hypothesized to explain the link between individuals' socioeconomic characteristics and their health behaviours, including physical activity, these factors pose serious measurement challenges. To complicate matters further, it is never clear to what extent these variables are mediating variables, as opposed to health outcomes themselves.

More recent research has come to realize that environmental factors both physical factors, such as neighbourhood sidewalks, and social factors, such as social support, are also critical for understanding people's physical activity behaviours because these factors provide the context in which people's activity decisions are made. This suggests that an unequal distribution of these environmental resources across socioeconomic groups may partially explain the socioeconomic gradient observed in physical activities. However, because individuals can, to some extent, control the environment they face, by for example choosing where they live, these environmental factors are ultimately still endogenous. Therefore, it is still premature to conclude that these factors are true causal determinants. Among the environmental factors studied in the literature, weather conditions are arguably the most unpredictable and therefore the most exogenous, in the sense people can not control the weather they face, at least once we control for the time of the year and the location in which individuals reside.

A recent study by Eisenberg and Okeke (2009) examines how environmental factors, especially the weather, affect physical activities and how this effect varies by socioeconomic status. Their results suggest that the effect of adverse weather conditions on physical activity is more pronounced among individuals in lower socioeconomic groups. The authors interpret this result as evidence of income constraints, which make it difficult for lower income individuals to substitute between indoor and outdoor physical activities. However, even if the weather is exogenous, socioeconomic status is not. This is particularly the case

[1] Using a questionnaire surveying a randomly selected sample of 2,298 Australian adults with questions on the preferred source of assistance or support to become physically active, barriers to regularly participate in physical activity, and preferred physical activities, they found significant variations between gender and among age groups.

when using education levels, as the authors do, to proxy socioeconomic status.

The difficulties of understanding why people in higher socioeconomic group participate more in physical activities not only come from the high endogeneity of the socioeconomic characteristics, but often also come from understanding the meaning of socioeconomic status. Socioeconomic status is conceptualized as the social standing or class of an individual or group. Although socioeconomic status can be measured by education, income or occupation, and these socioeconomic indicators are often strongly correlated, they in fact represent distinct aspects of a person's position in society and presumably exert their influences on people's behaviours through very different mechanisms. For example, education would be expected to capture knowledge-related assets, such as an appreciation of the importance of physical activity in determining health, as well as socio-psychological resource, such as self-efficacy and problem-solving coping capacity. Income, on the other hand, might be expected to primarily reflect the accessibility to material conditions affecting health behaviour, such as access to different forms of physical activity, that may provide very different levels of utility (or disutility). Therefore, the mechanisms of the effect of different social economic indicators should be studied separately.

This study focuses on the effect of income and attempts to establish a direct causal link between income and participation in physical activities. It does so by following the strategy of comparing the responsiveness of physical activity to weather fluctuations, but does so by focusing on more credible sources of variation in individual-level income that are more likely to reflect financial constraints. The basic idea is that having more physical activity options promotes participation in physical activities. This is particularly true when adverse weather deters outdoor activities and more costly indoor physical activities are the only viable option for higher income people to stay physically active. Therefore, if one of the reasons for higher income leading to more participation in physical activities is that it provides people with more options, the income effect ought to be larger under adverse weather conditions. Knowing whether or not this income effect on physical activities is causal is critical for designing effective policy interventions for promoting physical activities among low-income people. If there is no causal relation linking income to physical activities, then providing low-income people with more financial support, such as tax subsidies, would have no impact on their levels of physical activity. Moreover, given that the marginal increase in income may be spent very differently, this implies that if one of barriers low-income people face for participating in physical activity is indeed, as identified in this chapter, the lack of opportunities, directly providing them with easier access to exercise opportunities may be a more effective interventions for promoting physical activities in this social group.

A consumption choice model is first presented to formally demonstrate the mechanism of the income effect this study attempts to identify. In the model, when outdoor weather is nice, outdoor and indoor physical activities are assumed to be perfect substitutes in utility, that is they provide people with same utility gain, the value of which depends on the idiosyncratic preferred activity level embedded in an individual's personality. Outdoor options of physical activities are, however, preferred to indoor ones because they are relatively cheaper, so individuals whose preferred activity level is above some threshold will opt to participate in outdoor physical activities. In contrast, if weather conditions deter outdoor activities, such as when it rains or is extremely windy, there will be no utility gain from participating in outdoor physical activities, regardless of the preferred activity level of the person. In this situation, indoor activities are the only viable option. However, because of the higher cost of participating in indoor physical activities, people whose preferred activity level is not sufficiently high would not choose to participate in physical activities any more. In this scenario, the increase in income level plays a more important role in promoting participation in physical activities, compared with the situation when weather is nice and outdoor activity options are available. The model predicts that the income effect on participating in physical activities is greater under adverse weather, and the amplified effect is mainly through its promotion for indoor physical activities by making the indoor options more affordable.

To empirically test the income effect mechanism proposed in the model, correlations between income and physical activity participation are examined under different weather conditions. Using time-use data from three waves (1992, 1998 and 2005) of Statistics Canada's General Social Survey (GSS), a significant positive correlation between household income and physical activity participation is identified. More important, linking the time-use data to daily weather data from 56 Canadian cities, the findings show that the positive income-activity correlation becomes more pronounced as weather quality deteriorates. The results also show that the amplified income effect on participation in physical activities is mainly through promoting substitution from outdoor to indoor physical activity. Specifically, the effect of income on participation in indoor physical activities is enlarged more than 10 times when situation changes from having nice weather all day (no occurrence of precipitation or strong wind) to having rain or strong wind all day. Furthermore, we find this dependence of the income effect on the weather is robust to the inclusion of controls for individual self-reported health and when we instrument household income using a local unemployment rate. The findings appear consistent with the hypothesis that income effects are amplified under adverse weather by providing access to more costly alternative options.

The outline of the remainder of this chapter is as follows. The following section provides a more detailed review of the related literature. A theoretical model is then presented to demonstrate the mechanism this study attempts to capture. In section 4, the empirical methodology is described, including the data employed. Section 5 discusses the main results and the results of robustness check will be discussed in section 6, The chapter concludes with a section summarizing the main findings.

4.2 Related Literature

Promoting leisure time physical activities as a priority in improving public health has motivated numerous studies on this topic. Regardless of how socioeconomic status is measured, whether based on education, income or occupation, studies repeatedly find that people with lower socioeconomic status participate in less physical activity than their more advantageous counterparts. This is true for both the developed (Yen and Kaplan, 1998) and developing countries (Bauman, 2011); urban (Kavanagh et al., 2005) and rural areas (Parks, Housemann, and Brownson, 2003); adults (Craig, Russell, and Cameron, 2004) and youth (Humbert et al., 2006) and men (Azevedo et al., 2006) and women (Sternfeld, Ainsworth, and Quesenberry, 1999). Considering the wide range of populations these studies are based on, the evidence of the effects of socioeconomic status on physical activity is incontrovertible.

The underlying mechanisms driving these correlations are, however, unclear. One's social economics status is almost certainly correlated with unobservable individual characteristics, so studies only focusing on identifying individual factors affecting physical activity participation, which is the earlier studies tend to do, lack of capacity to establish a casual link between individuals' social economics status and their physical activity behaviour. A few theories in psychology, such as the theory of reasoned action (Ajzen and Fishbein, 1980) and the theory of planned behaviour (Ajzen, 1991), are used to explain the correlations. For example, the theory of reasoned action poses that an individual's behaviour involves a rational decision-making process and is performed under the individual's control, which requires the individual has all the necessary resources, skills and abilities to perform the behaviour at will. To apply this theory to explain physical activity behaviour, variations across socioeconomic groups could be due to differences in the capacity of making rational decisions and access to the resources needed. These theories propose numerous potential channels through which income or education could affect physical activities. They have not, however, been particularly helpful for precisely identifying any of these possibilities. Therefore, the association between socioeconomic status and physical activities remains unclear.

More recent studies have considered the effect of environmental factors on physical activities, and attribute variations in physical activity participation across socioeconomic groups to the interactive play of individual, social and physical environmental factors, which is often referred as the social ecological approach in literature. The strong appeal of this approach is reflected in the growing interest in the role of the environmental factors, especially physical environment attributes, in increasing physical activity levels. One of the reasons that can explain the heightened interest in environmental factors, is that they are highly policy relevant. For example, playgrounds and parks are levers that governments can easily manipulate in order to influence levels of physical activity in the population. The physical environment provides cues and opportunities for physical activities. Its influence is passive through the design of the urban environment, neighbourhood, domestic appliances and buildings which encourages or discourage incidental physical activities (Sallis and Owen, 1998). Meanwhile, studies also find that residence areas with high socioeconomic status, usually measured using area-level variables, such as unemployment rates, or median household incomes, have better access to sporting and recreational spaces (Powell, Slater, and Chaloupka, 2004); higher levels of safety and aesthetics (Wilson, Kirkland, Ainsworth and Addy, 2004); and even greater number of free-for-use physical activity resources (Estabrooks, Lee, and Gyurcsik, 2003). The influences of social environment include companionship, encouragement, and assistance from friends, family, and formal or informal social groups an individual belongs to (Lindstrom et al., 2001). Again, studies also show that these social factors, like the physical environmental factors, are also distributed across socioeconomic groups with favour to the higher ranked (Cerin and Leslie, 2008).

The social ecological approach, which uses environmental factors to explain the variations in physical activities across socioeconomic groups, suggests that the socioeconomic gradient in physical activity participation might be partially caused by the differences in the environments these groups face. However, what environment an individual faces can still be highly endogeneous. For example, sidewalks are less prevalent in low socioeconomic residency areas, and of course this can be part of the reasons why residents in these areas walk less. But, it can also be the case that people who choose to live in these areas are those who do not value physical activities to begin with. In this case, putting more sidewalks into these communities may have little or no impact on walking activity. Similarly, less group activities, such as those sports events organized by community, are less available to people with low socioeconomic status, but people who choose to socialize in those communities with less sports events are people who perhaps would not be interested in organized activities anyways. Therefore, using physical environment still can not isolate any particular underlying mechanism by effectively ruling out the confounding effect from individual unobserved characteristics.

The majority of the studies discussed above are from the health behaviour and behavioral medicine literatures. The common approach employed in these studies is to examine the correlates of physical activities. Economics, a discipline that studies how people allocate scarce resources in order to maximize their well-being, offers an alternative approach for studying people's physical activity behaviour. The theory of time allocation provides a theoretical foundation for understanding people's decisions to participate in physical activities. Developed from Becker's seminal work on time allocation (1965), the "SLOTH" framework proposed by Cawley (2004) offers a useful starting point for thinking about how people make decisions about physical activities. In the "SLOTH" framework people allocate their time to the following activities: sleep (S), leisure (including physical activity) (L), occupation (O), transportation (T), and home production (H). Thus, decisions regarding physical activities in this framework result from the utility-maximizing allocation of time, meaning the marginal utility of the last unit of time spent on each activity must be the same. Although the "SLOTH" framework is helpful for understanding rational people's physical activity behaviour and its underlying drivers, it fails to acknowledge that participating in physical activities requires not only time, but also money.

With the focus of studying the economic determinants of people's physical activity behaviour, Humphrey and Ruseski (2010) extend the "SLOTH" framework by acknowledging that people not only face time constraints but also financial constraints when making decision regarding their physical activities. Acknowledging the financial constraints people face when participating in physical activities is a very important step to take in modeling the effect of income on physical activities. They find that higher income is associated with a higher probability of participating in, but less time spent on, physical activities. Although distinguishing between the extensive and intensive margins when examining the income effect on physical activities makes progress towards establishing the causal links between income and physical activities, it is still difficult to interpret the correlations found in their study as causal relations due to the high endogeneity of income.

Eisenberg and Okeke (2009) also acknowledge that both time and financial constraints are affecting physical activity decisions, when they attempt to explain why the effects of exogenous weather fluctuations on physical activities vary with socioeconomic status. The effects of weather on physical activities have been well identified (see Tucker and Gilliland 2007 for a more complete review). However, to the best of my knowledge, Eisengerg and Okeke (2009) is the only study that explores how weather affects physical activities of different socioeconomic groups measured by either income or education differently. Considering many of the physical environmental factors, such as sidewalk facilities, have been used to provide insights into disparities in physical activity participation across socioeconomic groups, it is surprising that the effect of weather has not been much exploited. Linking the BRFSS data from 1993 to 2000 with weather data on a monthly level, Eisengerg and Okeke (2009) find within cold temperature ranges, a decrease in temperature causes a decrease in participation in physical activities and the effect is generally larger for people with lower income or lower education level.[2] Using exogenous weather variations is a superior approach, compared with other environmental factors which people can control, to avoid endogeneity issues in studying income effect. However, the shortcoming of using weather variations in this context is that weather may have direct effect on individual's actual health. Since one's health can in turn affect his or her physical activity behaviour (McNeil, Kreuter, and Subramanian, 2006), the effect of weather on health complicates the interpretation of the larger response to weather fluctuations observed in lower income

[2]The Behavioral Risk Factor Surveillance System (BRFSS) is a yearly telephone health survey system, tracking health conditions and risk behaviors in the United States.

groups. Moreover, the analysis in their study is conducted separately by income groups without controlling for the variations of education levels across these income groups. Since both an individual's income and physical activity behaviours are highly correlated with education, it is not clear what exactly the larger response to weather variations found in lower income groups captures.

4.3 Model

The model of physical activity behaviour in this chapter builds on the theoretical framework described in Humphreys and Ruseski (2010) by making a distinction between indoor and outdoor physical activities, where the utility of outdoor physical activities depends on weather conditions. The basic idea is that given time and budget constraints, individuals choose to allocate their time among physical activities, market work activities, and other activities, in order to maximize their utility, which is represented by a utility function comprising consumption of commodities and leisure:

$$\max_{t_o,t_i,t_z,C} U\{t_o,t_i,t_z,C\} = [I(\lambda)kt_o + kt_i + (1-k)t_z^{\alpha}]^{\eta}C^{1-\eta}, \quad 0<\alpha<1,\ 0<\eta<1 \tag{4.3.1}$$

where t_o, t_i, t_z denote time spent on outdoor physical activities, indoor physical activities, and other activities (including sedentary leisure activity, sleep, and home production activities), respectively, and C is the consumption of commodities. The utility gain from participating in physical activities, either indoors or outdoors, depends on the idiosyncratic preferred activity level embedded in the individual's personality, which is measure by k. k is assumed to be randomly and uniformly distributed on the interval $[0, 1]$ with 1 indicating the most active type of people. Utility associated with outdoor activities depends not only on a person's activity level k, but also on outdoor weather conditions λ. I is an indicator function, which takes the value 1 if outdoor weather quality is better than some threshold weather condition λ_0, and 0 otherwise. λ_0 can be considered as the weather quality that deters outdoor activities, specifically the incidence of precipitation and strong wind. This setting implies that when weather is nice outside, outdoor and indoor physical activities are equally enjoyable, and provide people with the same utility gain. However, if weather quality is worse than λ_0, outdoor physical activities will lose all its enjoyability, and do not contribute to people's well-being. In the real world, when people decide to exercise, they would choose either to do it outdoors or indoors, probably not both, especially within a one-day period. This reality is captured by the utility function where outdoor and indoor physical activities are assumed to be perfect substitutes. Moreover, the quasilinear feature of the utility function allows the possibility of individuals not spending any time on any physical activity as a result of optimal time allocation, which in fact closely resembles the real-world phenomenon that the majority people do not engage in any physical activity.

When making the optimal time allocation decision to maximize utility according to the utility function (4.3.1), an individual needs to consider two constraints: a financial constraint and a time constraint. The financial constraint is given by:

$$p_{em}t_o + F_i + p_{em}t_i + p_{zm}t_z + C = Y_0 + wh \tag{4.3.2}$$

where p_{em} is the direct cost associated with engaging in physical activities (indoors or outdoors), such as the cost of equipment maintenance or cost of hiring a personal trainer; p_{zm} is the direct cost associated with other activities; and the price of consumption is normalized to 1. An individual also needs to pay a fee F_i to have access to indoor facilities if he or she chooses to participate in indoor physical activities. In contrast, there is no charge for using outdoor resources for outdoor activities. Y_0 and w are a person's non-labour income and hourly wage rate, respectively. h denotes the time spent on market-work activity. This leads to the time constraint an individual faces, which is given by:

$$t_o + t_i + t_z + h = T \tag{4.3.3}$$

where T is the single-day time endowment. Combining the financial constraint (4.3.2) and time constraint (4.3.3), individuals choose t_o, t_i, t_z, and C to maximize their utility, subject to the full budget constraint:

$$p_e t_o + F_i + p_e t_i + p_z t_z + C = Y_0 + wT \tag{4.3.4}$$

where $p_e = p_{em} + w, p_z = p_{zm} + w$. p_e and p_z are the full prices associated with physical activities and other activities, taking into account the opportunity cost (wage).

Based on the utility function (4.3.1) and the full budget constraint (4.3.4), it is evident that outdoor physical activities are preferred to indoor ones when $\lambda > \lambda_0$, because outdoor physical activities provide the same marginal utility as indoor ones do without charging the access fee F_i. Since outdoor and indoor physical activities are perfect substitutes, in the situation when $\lambda > \lambda_0$, an individual maximizes the utility function $[kt_o + (1-k)t_z^\alpha]^\eta C^{1-\eta}$ subject to the constraint $p_e t_o + p_z t_z + C - Y_0 - wT = 0$, and the optimal decision can be obtained by solving the following Lagrangian problem:

$$L = [kt_o + (1-k)t_z^\alpha]^\eta C^{1-\eta} - \gamma(p_e t_o + p_z t_z + C - Y_0 - wT) \tag{4.3.5}$$

where γ is the Lagrange multiplier. The solution for optimal time allocated to outdoor physical activities is[3]:

$$t_o^* = \begin{cases} 0, & \text{if } k < \frac{(1-\eta+\alpha\eta)^{1-\alpha}(\alpha\eta)^\alpha p_e}{(1-\eta+\alpha\eta)^{1-\alpha}(\alpha\eta)^\alpha p_e + \eta p_z^\alpha (Y_0+wT)^{1-\alpha}} \\ \\ \frac{\eta(Y_0+T)}{p_e} - M, & \text{if } k \geq \frac{(1-\eta+\alpha\eta)^{1-\alpha}(\alpha\eta)^\alpha p_e}{(1-\eta+\alpha\eta)^{1-\alpha}(\alpha\eta)^\alpha p_e + \eta p_z^\alpha (Y_0+wT)^{1-\alpha}} \end{cases}$$

[3]Detailed solving procedures can be found in the Appendix.

where,

$$M = \frac{(1-k)^{\frac{1}{1-\alpha}}[\eta\alpha^{\frac{1}{1-\alpha}}p_e^{\frac{1}{1-\alpha}} + (1-\eta)\alpha^{\frac{\alpha}{1-\alpha}}p_e^{\frac{\alpha}{1-\alpha}}]}{k^{\frac{1}{1-\alpha}}p_z^{\frac{\alpha}{1-\alpha}}}. \tag{4.3.6}$$

The solution of t_o^* indicates that the probability of participating in outdoor physical activities is the probability that $t_o^* > 0$. Given a person's activity level, k, is randomly uniform distributed on the interval $[0,1]$, $(1-k)$ is the probability of an individual participating in physical activities. In the event that $\lambda > \lambda_0$, this probability can then be written:

$$Prob(LTPA|\lambda > \lambda_0) = \frac{\eta p_z^{\alpha}(Y_0 + wT)^{1-\alpha}}{(1-\eta+\alpha\eta)^{1-\alpha}(\alpha\eta)^{\alpha}p_e + \eta p_z^{\alpha}(Y_0+wT)^{1-\alpha}} \tag{4.3.7}$$

Equation (4.3.7) shows that a higher price of physical activities, p_e, negatively affects the probability of participating in physical activities, but a higher price of other activities, p_z, increases participation in physical activities. The effect of income Y_0 on participation in physical activities, which is of primary interest in this study, is derived from (4.3.7) in the following:

$$\frac{\partial Prob(LTPA|\lambda > \lambda_0)}{\partial Y_0} = \frac{\eta(1-\alpha)(\alpha\eta)^{\alpha}(1-\eta+\alpha\eta)^{1-\alpha}p_z^{\alpha}p_e(Y_0+wT)^{-\alpha}}{[(1-\eta+\alpha\eta)^{1-\alpha}(\alpha\eta)^{\alpha}p_e + \eta p_z^{\alpha}(Y_0+wT)^{1-\alpha}]^2} > 0 \tag{4.3.8}$$

Given both η and α are assumed to be fractions, the effect of Y_0 is unambiguously positive and the magnitude of the effect of Y_0 depends positively on the price of physical activities. The intuition is simply that the more expensive physical activities are, the more income is needed to gain the opportunity for participation.

Now, we turn to examine the effect of income on physical activities when weather is not nice and deters outdoor activities, that is where $\lambda < \lambda_0$. In this situation, outdoor physical activities contribute nothing to individuals' utility. Therefore, it must be indoor physical activities that an individual would consider to spend time on when he or she maximizes utility $[kt_i + (1-k)t_z^{\alpha}]^{\eta}C^{1-\eta}$ subject to the budget constraint $F_i + p_e t_i + p_z t_z + C - Y_0 - wT = 0$. When $\lambda < \lambda_0$, the optimal solution for time spent on indoor physical activities is then:

$$t_i^* = \begin{cases} 0, & \text{if } k < \frac{(1-\eta+\alpha\eta)^{1-\alpha}(\alpha\eta)^{\alpha}p_e}{(1-\eta+\alpha\eta)^{1-\alpha}(\alpha\eta)^{\alpha}p_e + \eta p_z^{\alpha}(Y_0+wT-F_i)^{1-\alpha}} \\ \\ \frac{\eta(Y_0+T-F_i)}{p_e} - M, & \text{if } k \geq \frac{(1-\eta+\alpha\eta)^{1-\alpha}(\alpha\eta)^{\alpha}p_e}{(1-\eta+\alpha\eta)^{1-\alpha}(\alpha\eta)^{\alpha}p_e + \eta p_z^{\alpha}(Y_0+wT-F_i)^{1-\alpha}}. \end{cases}$$

Again, given a person's activity level, k, is randomly and uniformly distributed on the interval $[0,1]$, the probability of an individual participating in physical activities is:

$$Prob(LTPA|\lambda < \lambda_0) = \frac{\eta p_z^{\alpha}(Y_0 + wT - F_i)^{1-\alpha}}{(1-\eta+\alpha\eta)^{1-\alpha}(\alpha\eta)^{\alpha}p_e + \eta p_z^{\alpha}(Y_0+wT-Fi)^{1-\alpha}} \tag{4.3.9}$$

and, the effect of income Y_0 is:

$$\frac{\partial Prob(LTPA|\lambda < \lambda_0)}{\partial Y_0} = \frac{\eta(1-\alpha)(\alpha\eta)^{\alpha}(1-\eta+\alpha\eta)^{1-\alpha} p_z^{\alpha} p_e (Y_0 + wT - F_i)^{-\alpha}}{[(1-\eta+\alpha\eta)^{1-\alpha}(\alpha\eta)^{\alpha} p_e + \eta p_z^{\alpha}(Y_0 + wT - Fi)^{1-\alpha}]^2} > 0. \quad (4.3.10)$$

From equation (4.3.7) and (4.3.9), we can see that conditional on the weather state, the probability of participating in physical activity when $\lambda > \lambda_0$, $Prob(LTPA|\lambda > \lambda_0)$, is necessarily greater than when $\lambda < \lambda_0$, $Prob(LTPA|\lambda < \lambda_0)$, meaning the probability of participating in physical activties is always higher when weather falls above some critical threshold. Comparing equation (4.3.8) and equation (4.3.10), non-labour income always has a positive effect on the probability of an individual participating in physical activities, regardless of the outdoor weather conditions. However, the effect in equation (4.3.10) is greater than in equation (4.3.8), which suggests income has a larger impact on physical activity participation under adverse weather conditions. The unambiguously positive effect of non-labour income on physical activities is expected because unlike income that comes from labour-market activity (wages), the increase in non-labour income does not change the opportunity cost of participating in physical activities. In other words, there is no confounding of the income effect by substitution effects between physical and market-work activities when analyzing the effect of Y_0.

Government policy or interventions promoting physical activities, such as subsidies for physical activity participation, can be evaluated within this model as equivalent to an increase in non-labour income. Taking a closer look at equations (4.3.8) and (4.3.10), the result $\frac{\partial Prob(LTPA|\lambda<\lambda_0)}{\partial Y_0} > \frac{\partial Prob(LTPA|\lambda>\lambda_0)}{\partial Y_0}$ is purely a consequence of the indoor activity access fee F_i. The intuition is that when F_i is required, physical activities become relatively more expensive. Therefore, the importance of income gets amplified by operating through one particular channel –providing easier access to more expensive physical activity options, such as indoor physical activity facilities. This mechanism is expected to be most pronounced on the extensive margin. Once people obtain the opportunity for participating in physical activities, income would probably still be expected to have a positive effect on the duration of engaging in physical activities as long as physical activities are believed to be normal goods. However, these is no reason to believe that the income effect on the duration of physical activities would be different under various weather conditions, which is also what the model predicts. Specifically from equations of t_o^* and t_i^*, we get:

$$\frac{\partial(t_o^*|\lambda > \lambda_0, t_o^* > 0)}{\partial Y_0} = \frac{\partial(t_i^*|\lambda < \lambda_0, t_i^* > 0)}{\partial Y_0} = \frac{\eta}{p_e} \quad (4.3.11)$$

where we can see that conditional on participation, the effect of income on the duration of physical activities, under any weather condition, depends on the price of physical activities and the relative importance of leisure time to consumption of commodities in the utility function.

4.4 Empirical Identification

To examine the real-world relevance of the mechanism identified in the theoretical model, the main objective in the empirical analysis is to test the following two hypotheses. First, the positive correlation between income and participation in physical activities (outdoor plus indoor) is larger when outdoor weather conditions prohibit opportunities for outdoor physical activities. Second, the amplified income effect on physical activity participation under adverse weather is mainly due to greater substitution from outdoor to indoor activities associated with higher income.

The tests of the hypotheses are conducted by linking weather data collected in 56 Canadian cities to the time-use data from three waves (1992, 1998, and 2005) of Statistics Canada's General Social Survey (GSS). The GSS time-use data identifies the detailed activities of survey respondents continuously over a randomly assigned 24-hour period. Compared with the monthly data, typical of health surveys, and used in the study by Eisenberg and Okeke (2009), GSS daily time-use data suits this study well in that it can be matched with weather information more precisely to capture the effect of weather on physical activities. Since the link is between the weather and activities on the same day, it is less likely that the correlation will be related to health effects of the weather, which is likely to happen with some lag. Restricting the analysis to individuals who are 15 years or older and have no long-term mental or physical health conditions that would limit their ability to participate in physical activities, the final sample includes 14,394 people with the information about their physical activities and the corresponding weather condition that they face between 6am and 11pm, an 18-hour time period during their reference day.

According to the detailed activity information in the GSS, 19 activities are identified as leisure time physical activities. For exploring the weather-induced substitution between outdoor and indoor physical activities, the activities need to be classified by whether they take place outdoors or indoors. Unfortunately, this information is not directly provided. For some activities such as hiking, golfing, and health club exercise, the outdoor/indoor classification is unambiguous, whereas for other activities, such as swimming, the classification is less clear. To deal with this shortcoming of the data, all physical activities are classified into three categories: outdoor, indoor, or ambiguous, where ambiguous indicates that the activities commonly take place either indoors or outdoors.

Table 4.1 shows the list of the leisure time physical activities and the classification and participation rates of each activity. A respondent would be considered as participating in one particular physical activities if he or she was recorded spending positive amount of time on that activity during the reference day. Among all physical activities, one of the outdoor activities – walking, hiking, jogging and running – has the highest participation rate (10.4%), followed by exercises, yoga, weight lifting (6.5%). In fact, these two activities together account for more than 60% of the incidences of people participating in physical activities. Conditional on the individual participating in physical activities during the reference day, 90.3% of them only participate in one of the three types of physical activities, either outdoor, or indoor, or ambiguous, and 9.7% participate more than one type. For those who participate in more than one type of physical activity, on average they spend more than 70% of their total physical activity time on only one type of the physical activity. Thus, according to the type of physical activities the respondent spent the most time on during the reference day, all people in our sample can be categorized into the following four groups: no physical activity; outdoor physical activity; indoor physical activity; ambiguous physical activity.[4]

The focus of this study is the effect of non-labour income on physical activities. Household income is used as the measure for respondent's non-labour income. In the GSS time-use questionnaire, household income is recorded in ranges: less than $10,000, $10,000 – $19,999, $20,000 – $29,999, $30,000 – $39,999, $40,000 – $49,999, $50,000 – $59,999, $60,000 – $79,999, and more than $80,000. Following Ruhm (2005), the level of income is coded as the midpoint of the range reported or 150% of the upper threshold. To correct for problem arising from comparing monetary income of households with different numbers of household members, the level of income is also adjusted by dividing the amount of income by the square root of family size as an equivalence scale, as is standard practice in the income distribution literature.

The weather data is taken from Environment Canada's National Climate Data and Information Archive (NCDIA). The weather information is collected for 56 Canadian cities on five weather elements – temperature, relative humidity, precipitation, wind speed, and cloud cover – at the top of each hour. In the theoretical model in section III, λ_0 defines a threshold in the distribution of weather, which distinguishes the weather conditions that are conducive for outdoor activities from those that are not. To operationalize this idea, weather conditions with the presence of precipitation or strong wind (wind speed in excess of 38km/hr) are defined as "adverse weather" in the sense that it prohibits outdoor activities.[5] Since the weather information in the NCDIA is collected at the top of each hour, the adverseness level of the daily weather is defined as the number of hours having "adverse weather" between 6am and 11pm.

The first hypothesis of concern – that the income effect on participation in overall physical activities is larger when the weather is worse – is tested using the following specification, which is estimated by probit regression:

$$\begin{aligned} Prob(LTPA_{icymd} = 1) =& \Phi \Big[\beta_0 + \beta_1 income_i + \beta_2 weather_{cymd} + \beta_3 income_i * weather_{cymd} \\ & + \theta_1 X_i + \theta_2 lfs_i + \delta_1 year_y + \delta_2 month_m + \delta_3 day_d + \delta_4 city_c \Big]. \quad (4.4.12) \end{aligned}$$

Subscripts denote, person i, city c, year y, month m, and reference day d, respectively. City and month fixed effects are included in order to limit the identification to weather variations that cannot be anticipated. The variable *day* is a vector of day of the week dummy variables, which is included to rule out the confounding effect that people with different income levels

[4]In the case where equal time are spent on two types of activities, the respondent is categorized by the type of the single activity he/she spent the most time on. Then, all individuals in the sample are able to be uniquely grouped into one of the three categories.

[5]The wind speed threshold corresponds to 8 or a "Strong Breeze" on the Beaufort Scale. At this speed: "Large tree branches are set in motion; whistling is heard in overhead wires; umbrella use becomes difficult; and empty plastic garbage cans tip over."

may have different schedules for other activities, such as work and family responsibility. X is a vector of demographic variables including education, age, gender, marital status, and the age of the person's youngest child. Although the effects of these factors on physical activities are unlikely to respond to weather fluctuations, which means they should not affect the estimate on the interaction term of income and weather, these factors are used in the analysis with the aim of improving the efficiency of the estimates. Lastly, a significant weakness of the GSS data is that it does not provide information on individual-level wage rates. In order to capture the variation in individuals' opportunity cost of participating in physical activities, the analysis also conditions on individual's labour force status lfs, which belongs to one of the following 6 groups: full-time workers; part-time workers; retirees; students; homemakers; and others (e.g. volunteers). Conditional on being employed, the respondent's industry and occupation information is also accounted for in the analysis. With this setup, the estimate of β_3 is what this study is most interested in. The estimate of this parameter is expected to be positive, meaning that the worse the weather conditions are, the more pronounced the positive income effect is.

As the mechanism identified in the model implies, income provides indoor activity options needed under adverse weather. Therefore, the second hypothesis states that the amplified income effect on physical activity participation, under adverse weather, mainly operates through its effect on indoor physical activities. Distinguishing outdoor, indoor, ambiguous physical activities, a multinomial logit model is then estimated. Specifically, the probability of participating in a particular type of activity is modeled as:

$$Prob(LTPA_{icymd} = j) = \frac{exp(\mu_{icymdj})}{1 + exp(\mu_{icymd1}) + exp(\mu_{icymd2}) + exp(\mu_{icymd3})} \tag{4.4.13}$$

where j=1,2,3 represent outdoor, indoor, and ambiguous physical activities, respectively; and μ_{icymd0} is normalized to zero, so that no participation in any physical activities during the reference day is the reference category. The linear index μ_{icymdj} are specified the same as equation (4.4.12), which is given by:

$$\begin{aligned} \mu_{icymdj} = \Big[& \beta_0 + \beta_1 income_i + \beta_2 weather_{cymd} + \beta_3 income_i * weather_{cymd} \\ & + \theta_1 X_i + \theta_2 lfs_i + \delta_1 year_y + \delta_2 month_m + \delta_3 day_d + \delta_4 city_c \Big] \end{aligned} \tag{4.4.14}$$

To support the hypothesis, we expect to see the probability of participating in indoor physical activities, relative to no physical activity, is more responsive to income under adverse weather, compared with outdoor physical activities. Therefore, the interaction term of income and weather is expected to be more positive in the case of indoor options, indicating that the income effect on overall physical activities becomes larger when weather gets worse mainly because its effect on indoor physical activities becomes larger.

4.5 Results

The sample mean values of the key variables used in the empirical analysis are shown in Table 4.2. On their reference day, 24.2% of the respondents participated in at least one type of physical activities. Comparing outdoor and indoor activities, outdoor physical activities are more popular than the indoor activities (12.2% compared to 6.3%). Between 6am to 11pm, 47.6% of the sample experienced at least one hour of adverse weather condition. Conditional on the occurrence of adverse weather, more than 50.0% of the time, the adverse weather lasted more than 4 hours; about 30.0% of the time lasted more than 8 hours; and 2.5% of the time it was present all day. The mean value of equivalent household income is $41,643, and the mean values of quintiles in the income distribution are $15,484, $25,803, $41,783, $62,402, and $110,300, respectively. There are 27.0% of the sample have a bachelor or graduate degree, while 30.7% have high school or less. 52.5% of the sample are male, 63.7% are married, and 57.4% have no kids. Full-time employees, retirees, and homemakers together account for 81.2% of the sample. The sample is rather evenly distributed across the months with the exception of December having fewer observations, which is due to that the GSS time-use survey is conducted for only half of the month in December.

For descriptive purposes, Table 4.3 presents the unconditional sample mean probabilities of physical activity participation across the key covariates. Comparing the rates of participation in physical activity across income quintiles, there is a clear tendency for physical activity to rise with income levels. From the first quintile to the fifth quintile of the income distribution, the participation rates rise monotonically from 21.8% to 27.5%, a 26.2% difference. Looking at the participation rates for outdoor, indoor, and ambiguous activities separately, indoor activities increase the most with income levels (37.9% difference) followed by outdoor activities (20.6% difference) and ambiguous activities (26.4%difference).

Among the three types of physical activities, the estimates in Table 4.3 suggest that outdoor physical activities are most negatively affected by adverse weather, as one would expect. Comparing a day when the weather is nice all day with a day when adverse weather conditions are present all day, the probability of participating in outdoor physical activities drops from 16.1% to 2.7%. In comparison, there is no clear patterns in changes of the incidence of indoor or ambiguous activities across weather qualities. The fact that outdoor physical activities are negatively affected by adverse weather may not be surprising itself, but towards what activities people substitute from outdoor activities and how the substitution depends on income levels are what is of primary interest in this chapter.

The differences in physical activity participation across education, gender, marital status, and age of the respondent's youngest child are all consistent with findings in existing literature. Single, more-educated and childless men tend to participate in physical activities more than their counterparts. A higher participation rate among people aged 65 and above, however, is a little surprising but may be partly explained by their lower opportunity cost of leisure time as non-labour market participants.

Across the 12 months, the unconditional rates of physical activity participation in Table 4.3 follow a clear seasonal pattern, with the highest values during the non-winter months (peaking in August at 32.2%) and the lowest values during the winter months (reaching a minimum of 16.6% in December). These patterns are also evident in Figure 4.1, which presents nonparametric regressions using a Kernel smoothing function. The monthly patterns, for both high- and low-income people, are mainly driven by fluctuations in outdoor activities, which also peak in summer months and drop significantly during winter months. But, the seasonality of participation in physical activity, including all three types of physical activities, is stronger among the low income people, which reflects the much lower participation during non-summer months and large increase in the summer. The monthly patterns of indoor activities are, in contrast, relatively flat. However, the increasing tendency is observed for indoor physical activities during winter months, most notably, among high income people. Seasonal variations in physical activities may be due to factors other than weather, such as changes in time allocation related to holidays, school, or work schedules. These results emphasize the importance of conditioning on the time of year when identifying the responsiveness of the correlation between income and physical activity to adverse weather shocks.

Before presenting the results of the regression analysis, the nonparametric regression analysis using a Kernel smoothing function is also conducted between income and physical activity participation, and the results are shown in Figure 4.2. Physical activity participation appears to increase with income regardless of the weather. However, the increase appears slightly stronger if adverse weather occurred at least one hour during the day. This appears particularly true when income lies above \$60,000. Regardless of the income levels, outdoor physical activity participation is significantly lower when there exist adverse weather conditions, and it only slightly increases with income. In contrast, in the presence of adverse weather conditions, there is a strong increasing trend in indoor activity with income. The comparison between outdoor and indoor activities under adverse weather implies that when weather is not conducive for outdoor activities, income starts to play a more important role in promoting physical activities through increasing indoor opportunities, suggesting substitution into indoor physical activities is more likely to happen among those with higher income.

Table 4.4 presents the results from estimating the baseline probit regression, given by equation (4.4.12), the effect of income on participating in physical activities, including outdoor, indoor, and ambiguous activities. The analysis starts with simpler specifications and arrives at equation (4.4.12), which is the goal of this study, by adding demographic controls and weather information. When the analysis is only conditional on year, month, day and city in specification (1), income appears to be strongly correlated with physical activity participation, which is consistent with findings in the existing literature. Specifically, \$10,000 increase in household income leads to 4.47% increase (from 0.2377 to 0.2439). Specification (2) adds detailed personal demographic controls. This serves to increase the income effect, so that at the mean values of the covariates, a \$10,000 increase in household income results in a 4.8% (from 0.2327 to 0.2439) increase in the probability of participating in a physical activity. When weather information is added in specification (3) ,although the adverse weather itself appears to have a significant negative impact on the probability of participating in physical activities, the effect of income is identical, which tells us that once we condition on city and month, the variation in weather is uncorrelated with income, as we would expect.

Specification (4) in Table 4.4, adds the interaction term of income and weather, which is the parameter of primary interest in testing the main hypothesis that the income effect on participating in physical activities is amplified by adverse weather. The results are consistent with the theory in that the significant and positive estimate of the interaction term indicates that the positive income effect becomes larger when weather condition becomes worse. This interaction effect is robust to excluding the education controls, as shown in Specification (5), although the effect of income itself increases in a non-trivial way. The fact that the estimate of the interaction term of income and weather is not sensitive to education is consistent with the theory that this particular underlying channel of income effect, which this study is trying to identify with weather fluctuations, is unlikely for other confounding factors to operate through. In fact, directly estimating the effects of the interaction terms between weather and other demographic variables besides income, such as education and gender, none of these interaction terms appears to have significant impact on physical activity participation. Based on the estimates from Specification (4), Table 4.5 presents the progressive changes in the effect of income on physical activities as the incidence of adverse weather in the day increases. In one extreme situation when there is no occurrence of adverse weather, the probability of participating in physical activities increases by 3.4% (from 0.2424 to 0.2508) when income increases by $10,000. However, in the other extreme situation when adverse weather is present all day, a $10,000 increase in income serves to increase participation in physical activities by 13.7% (from 0.1829 to 0.2080).

To further investigate the mechanism identified in the model, the second hypothesis is tested by estimating equation (4.4.13) to explore how weather fluctuations affect the income-activity correlations separately for indoor and outdoor physical activities. The results from this analysis are reported in Table 4.6. The main result is that the weather-income interaction term is only significant in predicting indoor physical activities. Although participation in outdoor physical activity increases with income, the magnitude of this relation does not vary with weather quality. The intuition behind the results is that when weather quality deteriorates, the correlation between indoor physical activity and income level is stronger probably because only people with higher income can afford to substitute from outdoor to indoor physical activity. However, when weather turns nice, because it is a cheaper option, every one, regardless of their income, would substitute from indoor to the outdoor options, which explains the reason why the interaction term of income and weather is insignificant in predicting outdoor physical activity.

Using the estimates in Table 4.6, Table 4.7 presents the predicted income effect (at the sample means) on the three types of physical activities, as the incidence of adverse weather increases in the day. As shown, higher income levels lead to higher probabilities of participating in all three types of physical activities (indoor, outdoor or ambiguous), relative to the probability of doing no physical activity (the odds ratio). When the incidence of adverse weather increases, the probability of participating in physical activities decreases, which is also true for all three types of activities. However, the increase in income mitigates this negative weather effect, where this income effect is most pronounced on indoor activities. When there is no adverse weather condition during the day, a $10,000 increase in income leads to a 2.8% increase in the relative probability of indoor physical activities (from 0.0523 to 0.0538). In comparison, when there exists adverse weather all day, a $10,000 increase in income leads to a 23.1% increase in the relative probability of indoor activities (from 0.0501 to 0.0616), which is more than 10 times larger than the income effect under nice weather. By contrast, the income effect on outdoor activities is not as responsive to weather conditions. When weather changes from being nice all day to raining all day, the effect of a \$10,000 increase in income only increases the odds ratio by 4.9% (from 0.1281 to 0.1343), compared to 6.9% (from 0.0578 to 0.0618), which is a moderate increase compared with the 10-times larger income effect on indoor activities. The findings strongly suggest that under adverse weather, the large increase in the effect of income on indoor activities accounts for the amplified income effect on participation in overall physical activities found in estimating equation (4.4.12).

Considering the intensive margin of people's physical activity, the mechanism of income effect proposed in this study – providing more activity options – implies that conditional on participation, the income effect on the duration of physical activities should not change with weather fluctuations significantly. This speculation about the income effect on the duration of physical activities is reasonable because the broader activity "options" only play an important role in encouraging people to participate by providing them with access to more activity opportunities, especially when adverse weather causes outdoor activities to be less enjoyable. Therefore, once people have already obtained access to physical activities, access to exercise opportunities is no longer an issue, so the particular mechanism of income effect – providing easier access to more options – is not important.

This speculation is predicted by the model in section III and supported by the results of regression analysis in Table 4.8. Conditional on participation, the effect of income on the duration of physical activities (minutes) is estimated, controlling for the same covariates as in previous analysis of the participation decisions. As shown, conditional on participating in physical activity, income level by itself, has marginally significant positive effect on the duration of physical activities. This positive effect of income on the duration might be due to that the high- and low-income people participate in very different physical activities, such as golfing versus walking, which require different amount of time to be devoted to. However, the magnitude of the income effect is not significantly affected by weather qualities.

4.6 Robustness Analysis

This study exploits exogenous weather variations to identify one particular mechanism through which income affects participation in physical activities. The advantage of using weather shocks is that the effects of other potential confounding factors influencing participation in physical activities are unlikely to change with weather, especially with the short-term fluctuations of interest here. Nonetheless, the endogeneity of income is still a potential concern for establishing a credible causal link. To confront this issue, the following two strategies are used to check the robustness of the findings.

First, given that there exists a strong correlation between income and health, and the weather is likely to have its own effect on people's health, the larger income effect on physical activities under adverse weather that we have identified may not be caused by the mechanism proposed in this study. Rather, it may reflect an effect of weather on health, since ill health has been identified in studies as an important barrier to physical activity (Chinn, et.al, 1992; McNeil, Kreuter, and Subramanian, 2006). In the context of using weather variations to understand the income effect on people's physical activity behaviour, health is arguably the most obvious and significant confounding factor. Fortunately, in the GSS data, respondents are asked to evaluate their own health level as excellent, good, fair, or poor. To obtain some evidence of this possible source of bias, this self-reported health is included as an additional control variable. Because the health information is only available in GSS time-use data in 1998 and 2005, the sample size of this robustness analysis is smaller than the one used in Table 4.4.

Table 4.9 presents the results of estimating equation (4.4.12) including the health control. With the subsample that contains health information, when health information is included in the analysis, the income effect is attenuated and becomes insignificant when there is no adverse weather conditions. However, the income effect is still significantly amplified by adverse weather, which is reflected in the same estimate of the interaction term as when health status is not controlled for. The results of this analysis imply that the evidence of the causal link between income and physical activities this study finds is robust to the potential confounding effects of health.

The second robustness check employs an instrumental variable IV estimator, addressing the endogeneity of income. Ideally, the instrument employed should have a strong correlation with individual income levels, but have no direct influence on individual decisions of whether or not to participate in physical activity. The average unemployment rate at the city level over the past twelve months is used to instrument income. The assumption is that the household income reported in the GSS data will be influenced by local labour market conditions reflected in the local unemployment rate, but conditional on the individual's labour force status, occupation and industry information, the local unemployment rate is uncorrelated with physical activities. The interaction term of unemployment rate and weather is used as the instrumental variable for the interaction term of income and weather.

The results of the estimation of the reduced form equation in the first stage are reported in Table 4.10. Most notably, the local unemployment rate appears to have a strong negative relation with income as expected, whereas the interaction term of income and weather is significantly correlated with the interaction of income and weather. All the instrumental variables together explain about 31.1% and 65.1% of the variations in the instrumented variables income and the interaction of income and weather, respectively. Therefore, the instrumental variables are strongly correlated with the potential endogenous variables, and the analysis do not appear to suffer from a weak instruments problem, which can severely bias the results in a finite sample.

Table 4.11 presents the results of IV probit estimation by maximum likelihood. As shown, the estimates of income as well as the interaction term of income and weather both remain the same as in the Specification (4) in Table 4.4, where no instrumental variables are employed. However, the magnitudes of the estimates become larger, as do the standard errors. Although the significance of the interaction term has dropped, which could be caused by the lost efficiency during the estimation process using IV, the positive estimate of the interaction term still indicates that there is no substantial changes in the results relative to the naive estimates. Moreover, based on the results of the Wald test for exogeneity, the exogeneity of the variables can not be rejected, which also provides us with some assurance of the reliability of findings in the main analysis.

As an additional robustness check, the same model is estimated using a different set of instrumental variables for household income. Specifically, examine the results when the analysis is limited to married people and the spouse's labour force status is used to instrument the household income of respondents. The spouse's labour force status is presumably correlated with his or her income level, which in turn has an impact on the household income level. Once other demographic characteristics are conditioned on, in particular the age of the youngest child in the household, the spouse's labour force status is plausibly exogenous to the respondent's own physical activity behaviour. According to the information available in the GSS, the labour force status of the spouse can be categorized into one of the following 6 groups: employed; looking for work; students; homemaker; retiree; and other. A set of dummy variables generated from this classification, together with the local unemployment rate, will be used as the instruments for household income. Their interaction terms with weather will serve as the instrumental variables for the interaction of income and weather.

The results from estimating the reduced form equation in the first stage are reported in Table 4.12. By limiting the sample to married respondents, about half of the observations are lost. However, the results indicate a strong correlation between the IVs and household income. Together the IVs account for 36.4% and 73.2% of the total variation in income and the interaction of income and weather, respectively.

The results of IV probit regression using spouses' labour force status as instruments are reported in Table 4.13. Because there are more instrumental than endogenous variables, an overidentification test of the validity of the instrumental variables is possible. Based on the Amemiya-Lee-Newey minimum chi-square statistics, the null hypothesis of valid instruments is not rejected (p=0.1696). The Wald test of exogeneity is marginal rejected, implying a potential endogeneity problem in the absence of IVs. However, the estimate of the interaction term of income and weather, again, remains positive, and the magnitude is comparable to the main estimates in Table 4.4 as well as to the result in Table 4.11 using the unemployment rate as the instrumental variable. Although the significance disappears, due to a larger standard error, which might be partly be caused by the drop of the sample size, the sign as well as the magnitude of the estimate of the interaction term is still consistent with the main results of this chapter.

4.7 Summary

This chapter proposes and empirically tests a mechanism by which income affects participation in physical activity. First, a theoretical model of the decision to participate in physical activities is developed, in which the indoor physical activity options are more expensive, but can provide higher utility gain than outdoor options when precipitation and/or strong wind occurs. It is demonstrated in the model that one mechanism of the income effect on physical activity participation is providing greater activity options, for example providing opportunities to substitute from outdoors to indoors when outdoor activities are not a viable option. Also, the model predicts that the positive correlation between income and physical activity participation becomes larger as the incidence of adverse weather during the day increases. The amplified income effect is mainly driven by a greater correlation between income and the indoor physical activity options, suggesting that the ability of individuals to substitute into a more conducive environment when outside weather deters physical activities is important for people's participation in physical activities.

To empirically test the mechanism, time-use data recording the daily detailed activities of individuals is linked with daily weather data. The variations in physical activity participation behaviours across income groups are examined using essentially random fluctuations in weather. The results point to a greater positive correlation between income and the probability of participating in physical activities when there is precipitation and/or strong wind deterring outdoor activities. Specifically, the estimates suggest that when there is no adverse weather during the day and outdoor activities are enjoyable, a \$10,000 increase in income leads to a 3.4% increase in the probability of participating in physical activities. However, the same increase in income leads to a 13.7% increase in the probability of participation in physical activities when rain and/or strong wind is present all day.

The results also show that the larger difference in physical activity behaviour across income groups under adverse weather is mainly driven by a greater difference in indoor activity participation. Under the two extreme weather conditions, the effects of a \$10,000 increase in income on the relative probability of indoor physical activities, compared with no physical activities, increases from 2.8% to 23.1%, while the effect of the \$10,000 increase on outdoor activities only changes from 4.9% to 6.9%.

These findings suggest that the negative effect of changes in environmental factors on physical activity participation can be alleviated by having more activity options, which often requires higher income. The physical activity participation decisions of the low-income population are more sensitive to changes in environmental factors. Therefore, the interventions and policies that increase physical activity options, such as providing easier access to indoor facilities, are likely to be effective for increasing the participation in physical activities, especially in the low-income population.

Table 4.1: Classification participation rate of leisure-time physical activities

Type of Leisure Time Physical Activities	Participation rates	Outdoor/Indoor/Ambiguous
Walking, hiking, jogging, running	10.37%	Outdoor
Exercises, yoga, weight lifting	6.47%	Indoor
Swimming, waterskiing	1.79%	Ambiguous
Football, basketball, baseball, volleyball, hockey, soccer, field hockey	1.62%	Ambiguous
Other outdoor activities/excursions	1.25%	Outdoor
Bicycling	1.04%	Outdoor
Golf	0.90%	Outdoor
Other sports, active leisure	0.85%	Outdoor
Bowling, pool, ping-pong, pinball	0.80%	Indoor
Skiing, ice skating, sledding, curling, snowboarding	0.63%	Ambiguous
Tennis, squash, racquetball, paddleball	0.54%	Ambiguous
Other sports (e.g. Frisbee, Catch)	0.45%	Ambiguous
Fishing	0.31%	Outdoor
Coaching	0.28%	Ambiguous
Boating	0.24%	Outdoor
Camping	0.17%	Outdoor
Judo, boxing, wrestling, fencing	0.15%	Indoor
Rowing, canoeing, kayaking, wind surfing, sailing	0.12%	Outdoor
Hunting	0.08%	Outdoor
Number of observations	14,394	

Notes: The participation rate is defined as the percentage of individuals who participate in the activity at any point in the day.

Source: Activity data from 1992, 1998 and 2005 General Social Survey (GSS), excluding respondents with activity limitations due to mental and/or physical health reasons.

Table 4.2: Means of key variables

	Mean
Dependent variables	
Participation in all leisure time physical activities	24.19%
Participation in outdoor leisure time physical activities	12.20%
Participation in indoor leisure time physical activities	6.46%
Participation in ambiguous leisure time physical activities	5.53%
Independent variables	
Weather (number of hours with precipitation and/or strong wind)	
0	52.42%
1	8.48%
2	5.73%
3	4.95%
4	4.07%
5	3.17%
6	2.74%
7	2.47%
8	2.06%
9	2.74%
10	2.11%
11	1.80%
12	1.51%
13	1.17%
14	0.92%
15	1.08%
16	0.69%
17	0.73%
18	1.16%
Household income ($, adjusted for household size)	41642.56
Education	
Graduate degree	7.30%
Bachelor degree	19.71%
Diploma	22.99%
Incomplete post-secondary education	19.35%
High school degree	15.23%
Less than high school	15.42%
Age of respondents	
15 – 24	15.52%
25 – 34	23.90%
35 – 44	24.09%
45 – 54	18.29%
55 – 64	9.81%
65+	8.39%

Male	52.50%
Married	63.66%
Age of respondents' youngest child	
No children	57.38%
0 – 4	14.35%
5 – 12	13.27%
13 – 18	8.75%
18+	6.25%
Months of the reference day	
January	7.73%
February	9.55%
March	9.08%
April	7.13%
May	8.50%
June	8.94%
July	9.19%
August	9.02%
September	8.09%
October	9.47%
November	8.97%
December	4.34%
Labour force status of respondents	
Full-time employee	61.06%
Part-time employee	5.43%
Retiree	9.55%
Student	8.79%
Homemaker	10.61%
Other	4.56%
Number of observations	14,394

Source: Activity data from 1992, 1998 and 2005 General Social Survey (GSS), excluding respondents with activity limitations due to mental and/or physical health reasons. Weather data from Canadian National Climate Data and Information Archive (NCDIA).

Table 4.3: Sample mean probabilities of participating in leisure time physical activities

	All	Leisure time physical activity		
		Outdoor	Indoor	Ambiguous
Weather				
(hours with precipitation and/or strong wind)				
0	0.2803	0.1709	0.0540	0.0554
1	0.2579	0.1435	0.0533	0.0611
2	0.2201	0.1152	0.0525	0.0525
3	0.2377	0.1080	0.0655	0.0642
4	0.2747	0.1405	0.0797	0.0545
5	0.2330	0.1301	0.0572	0.0456
6	0.2282	0.1011	0.0801	0.0470
7	0.1978	0.0904	0.0625	0.0448
8	0.1952	0.0606	0.0754	0.0593
9	0.1991	0.0800	0.0645	0.0545
10	0.2235	0.0891	0.0905	0.0439
11	0.1985	0.0576	0.1092	0.0317
12	0.2141	0.1010	0.0593	0.0537
13	0.1555	0.0523	0.0511	0.0521
14	0.2513	0.1037	0.0626	0.0850
15	0.2135	0.0883	0.0718	0.0534
16	0.2579	0.0281	0.1380	0.0918
17	0.1390	0.0378	0.0641	0.0372
18	0.1704	0.0267	0.0671	0.0766
Household income (adjusted by household size)				
Fifth quintile	0.2746	0.1366	0.0757	0.0623
Fourth quintile	0.2696	0.1357	0.0771	0.0569
Third quintile	0.2239	0.1105	0.0584	0.0550
Second quintile	0.2236	0.1138	0.0570	0.0528
First quintile	0.2176	0.1133	0.0549	0.0493
Education				
Graduate degree	0.3174	0.1927	0.0614	0.0633
Bachelor	0.2726	0.1304	0.0792	0.0630
Diploma	0.2244	0.1181	0.0621	0.0442
Incomplete post-secondary education	0.2342	0.1088	0.0694	0.0560
High school degree	0.1979	0.0987	0.0571	0.0421
Less than high school	0.2458	0.1231	0.0526	0.0701
Gender				
Male	0.2571	0.1232	0.0656	0.0683
Female	0.2250	0.1207	0.0635	0.0408

Marital status				
Married	0.2313	0.1304	0.0540	0.0469
Single	0.2605	0.1071	0.0834	0.0700
Age of respondents				
15 – 24	0.2609	0.0755	0.0904	0.0950
25 – 34	0.2357	0.1106	0.0695	0.0557
35 – 44	0.2120	0.1095	0.0451	0.0574
45 – 54	0.2247	0.1341	0.0532	0.0373
55 – 64	0.2568	0.1416	0.0731	0.0420
65+	0.3299	0.2267	0.0742	0.0289
Age of respondents' youngest child				
No child	0.2646	0.1293	0.0782	0.0571
0 – 4	0.2065	0.1150	0.0366	0.0550
5 – 12	0.2210	0.1106	0.0495	0.0609
13 – 18	0.2071	0.1146	0.0416	0.0509
18+	0.2068	0.1049	0.0688	0.0332
Month				
January	0.2184	0.0751	0.0634	0.0799
February	0.2320	0.0766	0.0999	0.0555
March	0.2180	0.0773	0.0847	0.0559
April	0.2398	0.1213	0.0748	0.0438
May	0.2226	0.1373	0.0477	0.0376
June	0.2863	0.1716	0.0551	0.0595
July	0.3087	0.1866	0.0429	0.0792
August	0.3224	0.2028	0.0550	0.0646
September	0.2458	0.1401	0.0631	0.0426
October	0.2110	0.1001	0.0584	0.0525
November	0.1886	0.0805	0.0626	0.0455
December	0.1657	0.0626	0.0696	0.0336
Number of observations	14,394			

Source: Activity data from 1992, 1998 and 2005 General Social Survey (GSS), excluding respondents with activity limitations due to mental and/or physical health reasons. Weather data from Canadian National Climate Data and Information Archive (NCDIA).

Table 4.4: Probit estimates of the probability of participating in leisure time physical activities

	Specification				
	(1)	(2)	(3)	(4)	(5)
Household income/1000	0.0034***	0.0036***	0.0036***	0.0026***	0.0032***
	(0.0006)	(0.0007)	(0.0007)	(0.0007)	(0.0007)
Weather	–	–	-0.0106***	-0.0262***	-0.0264***
			(0.0035)	(0.0067)	(0.0067)
(Household income/1000)*Weather	–	–	–	0.0004***	0.0004***
				(0.0001)	(0.0001)
Education					
Graduate degree	–	0.1765**	0.1826**	0.1809**	–
		(0.0732)	(0.0738)	(0.0737)	
Bachelor	–	0.0935*	0.0968*	0.0984*	–
		(0.0576)	(0.0581)	(0.0580)	
Diploma	–	-0.0137	-0.0122	-0.0010	–
		(0.0487)	(0.0497)	(0.0497)	
Incmplete post-secondary education	–	-0.0279	-0.0260	-0.0231	–
		(0.0499)	(0.0501)	(0.0501)	
High school	–	-0.1058*	-0.1026*	-0.1007*	–
		(0.0548)	(0.0549)	(0.0550)	
Age of respondents					
25 – 34	–	-0.0638	-0.0601	-0.0592	-0.0380
		(0.0569)	(0.0569)	(0.0571)	(0.0567)
35 – 44	–	-0.1529**	-0.1521**	-0.1506**	-0.1294**
		(0.0608)	(0.0608)	(0.0610)	(0.0610)
45 – 54	–	-0.1379**	-0.1366**	-0.1340**	-0.1152*
		(0.0656)	(0.0657)	(0.0659)	(0.0663)
55 – 64	–	-0.1421*	-0.1420*	-0.1361*	-0.1126
		(0.0752)	(0.0756)	(0.0755)	(0.0762)
65 and up	–	-0.1289	-0.1290	-0.1246	-0.1037
		(0.0832)	(0.0833)	(0.0830)	(0.0835)
Male	–	0.1370***	0.1381***	0.1371***	0.1486***
		(0.0312)	(0.0311)	(0.0312)	(0.0310)
Married	–	-0.0809**	-0.0811**	-0.0811**	-0.0792**
		(0.0333)	(0.0333)	(0.0332)	(0.0330)
Age of respondents' youngest child					
0 – 4	–	-0.0768*	-0.0782*	-0.0788	-0.0711
		(0.0477)	(0.0478)	(0.0478)	(0.0478)
13 – 18	–	-0.0207	-0.0205	-0.0194	-0.0254
		(0.0640)	(0.0640)	(0.0603)	(0.0642)

5 – 12	–	0.0288	0.0302	0.0309	0.0350
		(0.0486)	(0.0488)	(0.0487)	(0.0487)
19 and up	–	-0.0993	-0.1011	-0.0999	-0.1043
		(0.0727)	(0.0728)	(0.0727)	(0.0730)
Labour Force Status fixed effects	–	yes	yes	yes	yes
Occupation fixed effects	–	yes	yes	yes	yes
Industry fixed effects	–	yes	yes	yes	yes
Day of week fixed effects	yes	yes	yes	yes	yes
Month fixed effects	yes	yes	yes	yes	yes
Year fixed effects	yes	yes	yes	yes	yes
City fixed effects	yes	yes	yes	yes	yes
Pseudo R^2	0.0199	0.0410	0.0419	0.0427	0.0406
Number of observations	14,366	14,366	14,366	14,366	14,366

Notes: Standard errors are clustered by city and time (year, month, and day). The weather is measured as the number of hours with precipitation and/or strong wind between 6am to 11pm during the reference day. 28 observations are dropped as they predict failure perfectly. ***,**,* indicate statistical significance at the 1%, 5%, and 10% levels, respectively.
Source: Activity data from 1992, 1998 and 2005 General Social Survey (GSS), excluding respondents with activity limitations due to mental and/or physical health reasons. Weather data from Canadian National Climate Data and Information Archive (NCDIA).

Table 4.5: Predicted probabilities of participating in leisure time physical activities, by weather

	Probability of participating in Leisure Time Physical Activities		
	Mean income	Mean income+$10,000	% increase
Weather (number of hours with precipitation and/or strong wind)			
(hours with adverse weather)			
0	24.24%	25.08%	3.44%
1	23.89%	24.83%	3.94%
2	13.54%	24.58%	4.44%
3	23.19%	24.34%	4.95%
4	22.84%	24.09%	5.48%
5	22.49%	23.85%	6.00%
6	22.15%	23.60%	6.54%
7	21.81%	23.36%	7.09%
8	21.48%	23.12%	7.64%
9	21.15%	22.88%	8.21%
10	20.82%	22.64%	8.78%
11	20.49%	22.41%	9.36%
12	20.17%	22.17%	9.95%
13	19.85%	21.94%	10.55%
14	19.53%	21.71%	11.16%
15	19.22%	21.48%	11.77%
16	18.91%	21.25%	12.40%
17	18.60%	21.02%	13.04%
18	18.29%	20.80%	13.69%

Notes: The probabilities of participating in leisure time physical activities at each weather condition are predicted at the mean values of the sample.
Source: The mean values of the variables used for prediction are from 1992, 1998 and 2005 General Social Survey (GSS), excluding respondents with activity limitations due to mental and/or physical health reasons.

Table 4.6: Multinomial Logit estimates of the probability of participating in leisure time physical activities

	Outdoor	Indoor	Ambiguous
Household income/1000	0.0047***	0.0027	0.0061***
	(0.0016)	(0.0022)	(0.0022)
Weather	-0.0488***	-0.0442**	-0.0314
	(0.0155)	(0.0181)	(0.0205)
(Household income/1000)*Weather	0.0001	0.0010***	0.0006
	(0.0003)	(0.0003)	(0.0004)
Education			
Graduate degree	0.4854***	0.2743	0.0165
	(0.1515)	(0.2214)	(0.2202)
Bachelor	0.1498	0.4583**	-0.0597
	(0.1242)	(0.1804)	(0.0574)
Diploma	0.0373	0.2228	-0.3959**
	(0.1139)	(0.1742)	(0.1597)
Incomplete post-secondary education	0.0494	0.2055	-0.4060**
	(0.1191)	(0.1731)	(0.1709)
High school	-0.1608	0.0766	-0.4715***
	(0.1244)	(0.1876)	(0.1709)
Age of respondents			
25 – 34	0.1654	-0.1989	-0.3033*
	(0.1499)	(0.1620)	(0.1624)
35 – 44	0.1347	-0.6509**	-0.4264**
	(0.1588)	(0.1780)	(0.1799)
45 – 54	0.3279**	-0.5492***	-0.8913***
	(0.1615)	(0.1854)	(0.2160)
55 – 64	0.0961	-0.2588	-0.7971***
	(0.1799)	(0.2184)	(0.2415)
65 and up	0.2864	-0.2279	-1.4186***
	(0.1971)	(0.2468)	(0.3384)
Male	0.1572**	0.0527	0.6029***
	(0.0739)	(0.1008)	(0.1127)
Married	0.0382	-0.2831***	-0.3255***
	(0.0776)	(0.0994)	(0.1136)
Age of respondents' youngest child			
0 – 4	-0.1002	-0.5459***	0.2740 *
	(0.1162)	(0.1682)	(0.1676)
5 – 12	-0.0861	-0.0469	0.5405***
	(0.1132)	(0.1504)	(0.1653)

13 – 18	-0.1140	-0.2524	0.4906**
	(0.1488)	(0.2076)	(0.2211)
19 and up	-0.3649**	0.0515	0.0293
	(0.1578)	(0.2168)	(0.2967)
Labour Force Status fixed effects	yes	yes	yes
Occupation fixed effects	yes	yes	yes
Industry fixed effects	yes	yes	yes
Day of week fixed effects	yes	yes	yes
Month fixed effects	yes	yes	yes
City fixed effects	yes	yes	yes
Pseudo R^2		0.0782	
Number of observations		14,385	

Notes: Standard errors are clustered by city and time (year, month, and day). The weather is measured as the number of hours with precipitation and/or strong wind between 6am to 11pm during the reference day. ***,**,* indicate statistical significance at the 1%, 5%, and 10% levels, respectively. **Source:** Activity data from 1992, 1998 and 2005 General Social Survey (GSS), excluding respondents with activity limitations due to mental and/or physical health reasons. Weather data from Canadian National Climate Data and Information Archive (NCDIA).

Table 4.7: Predicted probabilities of leisure time physical activities (odds ratios), by weather and activity type

	Outdoor			Indoor			Ambiguous		
	Mean (income)	Mean+$10,000 (income)	%increase	Mean (income)	Mean+$10,000 (income)	%increase	Mean (income)	Mean+$10,000 (income)	%increase
Adverseness weather									
0	0.1281	0.1343	4.87%	0.0523	0.0538	2.75%	0.0377	0.0401	6.32%
1	0.1226	0.1287	4.98%	0.0522	0.0542	3.78%	0.0375	0.0401	6.98%
2	0.1173	0.1232	5.09%	0.0521	0.0546	4.83%	0.0373	0.0401	7.64%
3	0.1122	0.1180	5.21%	0.0519	0.0550	5.89%	0.0371	0.0401	8.31%
4	0.1073	0.1131	5.32%	0.0518	0.0554	6.95%	0.0369	0.0402	8.98%
5	0.1027	0.1083	5.44%	0.0517	0.0558	8.03%	0.0366	0.0402	9.65%
6	0.0982	0.1037	5.55%	0.0516	0.0563	9.12%	0.0364	0.0402	10.33%
7	0.0940	0.0993	5.67%	0.0514	0.0567	10.22%	0.0362	0.0402	11.01%
8	0.0899	0.0951	5.78%	0.0513	0.0571	11.33%	0.0360	0.0402	11.70%
9	0.0860	0.0911	5.90%	0.0512	0.0576	12.46%	0.0358	0.0402	12.39%
10	0.1281	0.1343	4.87%	0.0511	0.0580	13.59%	0.0356	0.0403	13.09%
11	0.0788	0.0836	6.13%	0.0509	0.0584	14.73%	0.0354	0.0403	13.78%
12	0.0753	0.0801	6.25%	0.0508	0.0589	15.89%	0.0352	0.0403	14.49%
13	0.0721	0.0767	6.36%	0.0507	0.0593	17.06%	0.0350	0.0403	15.20%
14	0.0690	0.0734	6.48%	0.0506	0.0598	18.24%	0.0348	0.0403	15.91%
15	0.0660	0.0703	6.59%	0.0504	0.0602	19.43%	0.0346	0.0403	16.63%
16	0.0631	0.0674	6.71%	0.0503	0.0607	20.64%	0.0344	0.0404	17.35%
17	0.0604	0.0645	6.82%	0.0502	0.0612	21.85%	0.0342	0.0404	18.07%
18	0.0578	0.0618	6.94%	0.0501	0.0616	23.08%	0.0340	0.0404	18.80%

Notes: The probabilities of participating in leisure time physical activities at each weather condition, relative to no participation in physical activities, are predicted at the mean values of the sample. The adverse weather is the number of hours there exists precipitation and/or strong wind during the reference day.
Source: The mean values of the variables used for prediction are from 1992, 1998 and 2005 General Social Survey (GSS), excluding respondents with activity limitations due to mental and/or physical health reasons.

Table 4.8: OLS estimates of the duration of leisure time physical activities, conditional on participation

	Coefficient	Standard error
Household income/1000	0.2132*	0.1090
Weather	0.4234	0.8361
(Household income/1000)*Weather	-0.0005	0.0155
Education		
Graduate degree	-2.6364	8.5894
Bachelor	-8.6389	6.7810
Diploma	-1.7372	6.7478
Incomplete post-secondary education	-6.1177	6.3735
High school	2.2549	7.5660
Age of respondents		
25 – 34	-6.5198	6.3735
35 – 44	-4.5369	7.9694
45 – 54	-24.3435**	8.2044
55 – 64	-15.3699*	8.8642
65 and up	-14.9015	10.2123
Male	16.4277***	4.1242
Married	-9.7273**	4.3141
Age of respondents' youngest child		
0 – 4	3.2469	7.6123
5 – 12	3.5975	6.7155
13 – 18	16.2764**	8.6130
19 and up	-7.8737	7.4222
Days of the week		
Monday	-41.5592***	6.7014
Tuesday	-44.5837***	6.5221
Wednesday	-40.1849***	6.7077
Thursday	-44.5837***	6.6755
Friday	-47.0213***	6.9704
Saturday	3.2415***	8.7806
Labour Force Status fixed effects	yes	–
Occupation fixed effects	yes	–
Industry fixed effects	yes	–
Month fixed effects	yes	–
City fixed effects	yes	–
Pseudo R^2	0.1435	
Number of observations	3,528	

Notes: Standard errors are clustered by city and time (year, month, and day). The weather is measured as the number of hours with precipitation and/or strong wind between 6am to 11pm during the reference day. ***,** ,* indicate statistical significance at the 1%, 5%, and 10% levels, respectively.
Source: Activity data from 1992, 1998 and 2005 General Social Survey (GSS), excluding respondents with activity limitations due to mental and/or physical health reasons. Only respondents whose reported duration of physical activities greater than 0 are included in regression. Weather data from Canadian National Climate Data and Information Archive (NCDIA).

Table 4.9: Robustness analysis using health status as an additional control variable

	Without health control	With Health Control
Household income/1000	0.0019**	0.0014
	(0.0008)	(0.0009)
Weather	-0.0296***	-0.0314***
	(0.0083)	(0.0082)
(Household income/1000)*Weather	00004**	0.0004**
	(0.0002)	(0.0002)
Education		
Graduate degree	0.2110**	0.1754**
	(0.0814)	(0.0826)
Bachelor	0.1628**	0.1353**
	(0.0634)	(0.0637)
Diploma	0.0558*	0.0420
	(0.0611)	(0.0612)
Incomplete post-secondary education	0.0920	0.0734
	(0.0581)	(0.0585)
High school	-0.0551	-0.0656
	(0.0658)	(0.0655)
Age of respondents		
25 – 34	-0.0707	-0.0409
	(0.0702)	(0.0707)
35 – 44	-0.1821**	-0.1507**
	(0.0746)	(0.0759)
45 – 54	-0.1533**	-0.1300*
	(0.0764)	(0.0758)
55 – 64	-0.0906	-0.0741
	(0.0872)	(0.0888)
65 and up	-0.0701	-0.0157
	(0.0999)	(0.0989)
Male	0.1243***	0.1253***
	(0.0356)	(0.0364)
Married	-0.0793**	-0.0858**
	(0.0374)	(0.0381)
Age of respondents' youngest child		
0 – 4	-0.0487	-0.0646
	(0.0583)	(0.0585)
5 – 12	0.0867	0.0726
	(0.0563)	(0.0573)
13 – 18	0.0096	-0.0042
	(0.0720)	(0.0714)
19 and up	-0.0579	-0.0659
	(0.0789)	(0.0785)

Health Status		
Excellent	–	0.4600***
		(0.0713)
Good	–	0.2185***
		(0.0689)
Fair	–	0.0395
		(0.0695)
Labour Force Status fixed effects	yes	yes
Occupation fixed effects	yes	yes
Industry fixed effects	yes	yes
Day of week fixed effects	yes	yes
Month fixed effects	yes	yes
Year fixed effects	yes	yes
City fixed effects	yes	yes
Pseudo R^2	0.0422	0.0544
Number of observations	10,868	10,868

Notes: Standard errors are clustered by city and time (year, month, and day). The weather is measured as the number of hours with precipitation and/or strong wind between 6am to 11pm during the reference day. ***,**,* indicate statistical significance at the 1%, 5%, and 10% levels, respectively.
Source: Activity data from 1998 and 2005 General Social Survey (GSS), excluding respondents with activity limitations due to mental and/or physical health reasons. Weather data from Canadian National Climate Data and Information Archive (NCDIA).

Table 4.10: OLS estimates of the first stage reduced form equation (unemployment rate as IV)

	Endogenous variables			
	Household income/1000		(Household income/1000)*weather	
	Estimates	Standard Error	Estimates	Standard Error
Unemployment	-67.1232***	24.4581	189.1983	153.9865
Weather	-0.0755	0.1431	51.0592***	0.9007
Unemployment*Weather	0.8704	1.6681	-122.9453***	10.5024
Education				
Graduate degree	23.2562***	0.9202	73.0286***	5.7938
Bachelor	15.1667***	0.7200	40.8319***	4.5331
Diploma	7.2567***	0.6646	17.5613***	4.1846
Incomplete post-secondary education	5.9134***	0.6864	10.9429**	4.3217
High school	4.2921***	0.7160	8.0105*	4.5076
Age of respondents				
25 – 34	0.8779	0.7527	0.3947	4.7389
35 – 44	4.6098***	0.8016	7.8297	5.0470
45 – 54	7.6965**	0.8403	19.1314***	5.2902
55 – 64	5.7964***	0.9394	0.9607	5.9147
65 and up	-0.0312	1.1292	-14.3512**	7.1097
Male	3.6374***	0.4312	9.9436***	2.7150
Married	8.5691***	0.4386	23.7586***	2.7614
Age of respondents' youngest child				
0 – 4	-13.5825***	0.6657	-39.3115***	4.1915
5 – 12	-13.8528***	0.6562	-39.4701***	4.1313
13 – 18	-11.9931***	0.8573	-36.4932***	5.3973
19 and up	-5.2085***	0.9706	-18.0653***	6.1106
Labour Force Status fixed effects	yes	–	yes	–
Occupation fixed effects	yes	–	yes	–
Industry fixed effects	yes	–	yes	–
Day of week fixed effects	yes	–	yes	–
Month fixed effects	yes	–	yes	–
Year fixed effects	yes	–	yes	–
City fixed effects	yes	–	yes	–
R^2	0.3110		0.6599	
Prob>F	0.0000		0.0000	
Number of observations	13,915		13,915	

Notes: Standard errors are clustered by city and time (year, month, and day). The weather is measured as the number of hours with precipitation and/or strong wind between 6am to 11pm during the reference day. ***,** ,* indicate statistical significance at the 1%, 5%, and 10% levels, respectively.
Source: Income data from 1992, 1998 and 2005 General Social Survey (GSS), excluding respondents with activity limitations due to mental and/or physical health reasons. Weather data from Canadian National Climate Data and Information Archive (NCDIA).

Table 4.11: Robustness analysis using instrumental variables (unemployment rate) probit estimator

	Estimates	Standard Error
Household income/1000	0.0123	0.0234
Weather	-0.0572	0.0381
(Household income/1000)*Weather	0.0011	0.0009
Education		
Graduate degree	-0.0853	0.5562
Bachelor	-0.0310	0.3635
Diploma	-0.0723	0.1781
Incomplete post-secondary education	-0.0303	0.1474
High school	-0.1062	0.1125
Age of respondents		
25 – 34	-0.1121**	0.0527
35 – 44	-0.2110*	0.1213
45 – 54	-0.2381	0.1914
55 – 64	-0.2283	0.1490
65 and up	-0.1602**	0.0726
Male	0.0952	0.0906
Married	-0.1617	0.2052
Age of respondents' youngest child		
0 – 4	0.0980	0.3254
5 – 12	0.2415	0.3322
13 – 18	0.1006	0.2912
19 and up	-0.0145	0.1402
Labour Force Status fixed effects	yes	–
Occupation fixed effects	yes	–
Industry fixed effects	yes	–
Day of week fixed effects	yes	–
Month fixed effects	yes	–
Year fixed effects	yes	–
City fixed effects	yes	–
Wald chi2 (114)	510.03	
Number of observations	13,903	
Wald test of exogeneity	$chi2(2) = 1.08$	$Prob > chi2 = 0.5281$

Notes: Standard errors are clustered by city and time (year, month, and day). The weather is measured as the number of hours with precipitation and/or strong wind between 6am to 11pm during the reference day. ***,** ,* indicate statistical significance at the 1%, 5%, and 10% levels, respectively.
Source: Activity data from 1992, 1998 and 2005 General Social Survey (GSS), excluding respondents with activity limitations due to mental and/or physical health reasons. Weather data from Canadian National Climate Data and Information Archive (NCDIA).

Table 4.12: OLS estimates of the first stage reduced form equation (spouse's labour force status and unemployment rate as IVs)

	Endogenous variables			
	Household income/1000		(Household income/1000)*weather	
	Estimates	Standard Error	Estimates	Standard Error
Unemployment	-90.2573***	29.5532	163.2677	193.2961
Unemployment*Weather	4.1171**	1.9999	-101.3655***	27.5010
Spouses' labour force status				
Employed	12.5082***	1.7511	2.4099	6.6666
Looking for job	1.1194	21.2133	14.7462*	8.7583
Student	-2.6224	2.3644	25.4336***	8.6857
Homemaker	0.5622	1.9309	1.2929	7.4333
Retiree	-0.0560	2.1109	7.5374	8.5666
Employed*Weather	-0.0740	0.1670	11.7582***	2.0652
Looking for job*Weather	-0.1220	0.3469	-4.3594	4.2818
Student*Weather	-0.2869	0.2456	-13.3670***	2.6448
Homemaker*Weather	-0.1256	0.2039	-0.2854	2.5252
Retiree*Weather	0.1118	0.3092	0.8611	3.8739
Weather	-0.2502	0.2402	45.6643***	3.1165
Education				
Graduate degree	23.9735***	1.3096	77.8881***	9.3754
Bachelor	17.9761***	0.9396	47.9520***	5.7529
Diploma	9.2643***	0.8014	22.5567***	4.9235
Incomplete post-secondary education	8.8372***	0.8845	21.0208***	5.5322
High school	5.9973***	0.7902	12.2145**	4.9496
Age of respondents				
25 – 34	10.1357***	1.0948	26.0878***	7.0805
35 – 44	14.2317***	1.1362	38.3195***	7.1896
45 – 54	18.6995***	1.2188	53.7227***	7.8103
55 – 64	19.0688***	1.4770	41.8755***	8.6803
65 and up	14.5676***	1.8172	23.9492**	11.2971
Male	2.4650***	0.5954	5.0222	3.7794
Age of respondents' youngest child				
0 – 4	-11.5978***	0.7816	-28.9493***	4.7627
5 – 12	-13.9355***	0.8290	-39.2298***	5.1658
13 – 18	-12.9459***	1.0103	-35.9484***	6.0507
19 and up	-6.4945***	1.0475	-21.8431***	6.8459

Labour Force Status fixed effects	yes	–	yes	–
Occupation fixed effects	yes	–	yes	–
Industry fixed effects	yes	–	yes	–
Day of week fixed effects	yes	–	yes	–
Month fixed effects	yes	–	yes	–
Year fixed effects	yes	–	yes	–
City fixed effects	yes	–	yes	–
R^2	0.3639		0.7325	
Prob>F	0.0000		0.0000	
Number of observations	7,642		7,642	

Notes: Standard errors are clustered by city and time (year, month, and day). The weather is measured as the number of hours with precipitation and/or strong wind between 6am to 11pm during the reference day. ***,**,* indicate statistical significance at the 1%, 5%, and 10% levels, respectively.
Source: Income data from 1992, 1998 and 2005 General Social Survey (GSS), excluding respondents with activity limitations due to mental and/or physical health reasons. Weather data from Canadian National Climate Data and Information Archive (NCDIA).

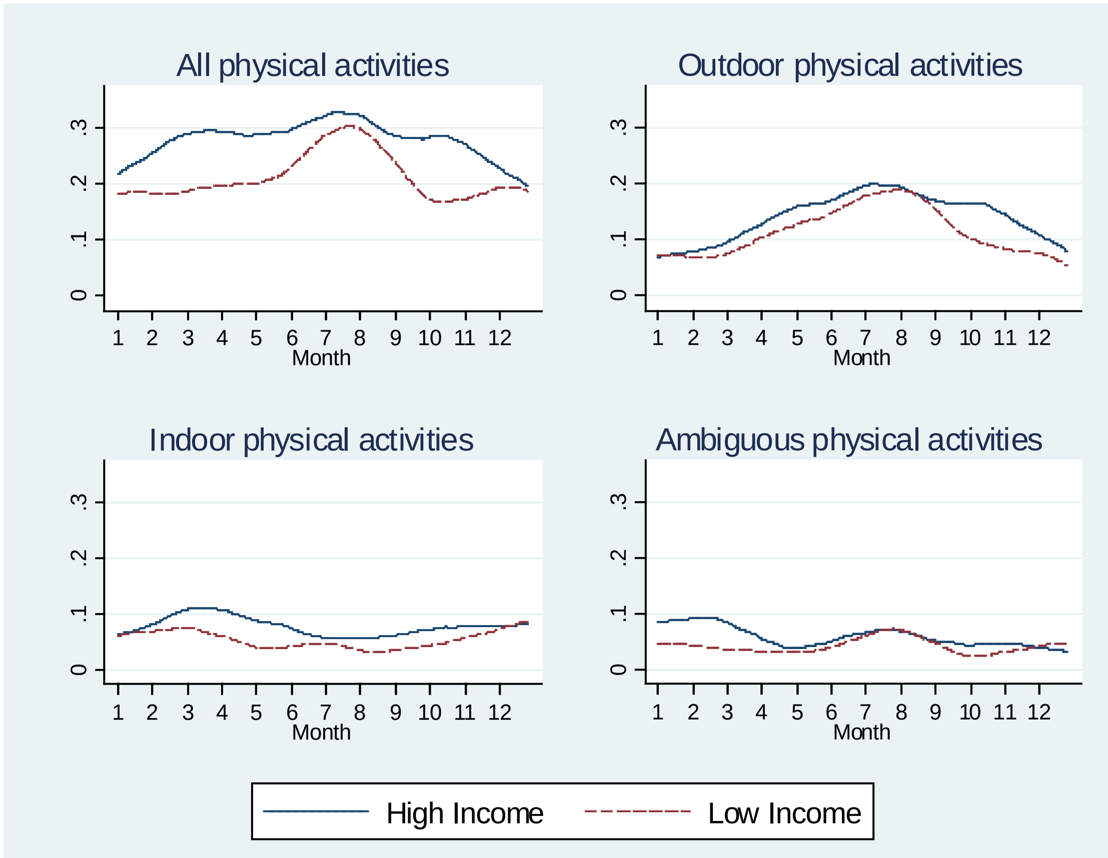
Table 4.13: Robustness analysis using instrumental variables (spouse's labour force status and unemployment rate as IVs) probit estimator

	Estimates	Standard Error
Household income/1000	-0.0053	0.0035
Weather	-0.0431*	0.0235
(Household income/1000)*Weather	0.0008	0.0005
Education		
Graduate degree	0.4090***	0.1110
Bachelor	0.2986***	0.0883
Diploma	0.1026	0.0692
Incomplete post-secondary education	0.0961	0.0735
High school	0.0220	0.0696
Age of respondents		
25 – 34	-0.0281	0.1024
35 – 44	-0.0641	0.1094
45 – 54	0.0190	0.1172
55 – 64	-0.0241	0.1194
65 and up	0.0224	0.1249
Male	0.1045***	0.0393
Age of respondents' youngest child		
0 – 4	-0.1438**	0.0670
5 – 12	0.0163	0.0677
13 – 18	-0.1674**	0.0776
19 and up	-0.1280*	0.0774
Labour Force Status fixed effects	yes	–
Occupation fixed effects	yes	–
Industry fixed effects	yes	–
Day of week fixed effects	yes	–
Month fixed effects	yes	–
Year fixed effects	yes	–
City fixed effects	yes	–
Wald chi2 (114)	376.16	
Number of observations	7,633	
Test of overidentifying restrictions:		
Amemiya-Lee-Newey minimun chi-sq statistic:	14.075 Chi-sq (10)	P-value: 0.1696
Wald test of exogeneity	$chi2(2) = 4.81$	$Prob > chi2 = 0.0904$

Notes: Standard errors are clustered by city and time (year, month, day). The weather is measured as the number of hours with precipitation and/or strong wind between 6am to 11pm during the reference day. ***,** ,* indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

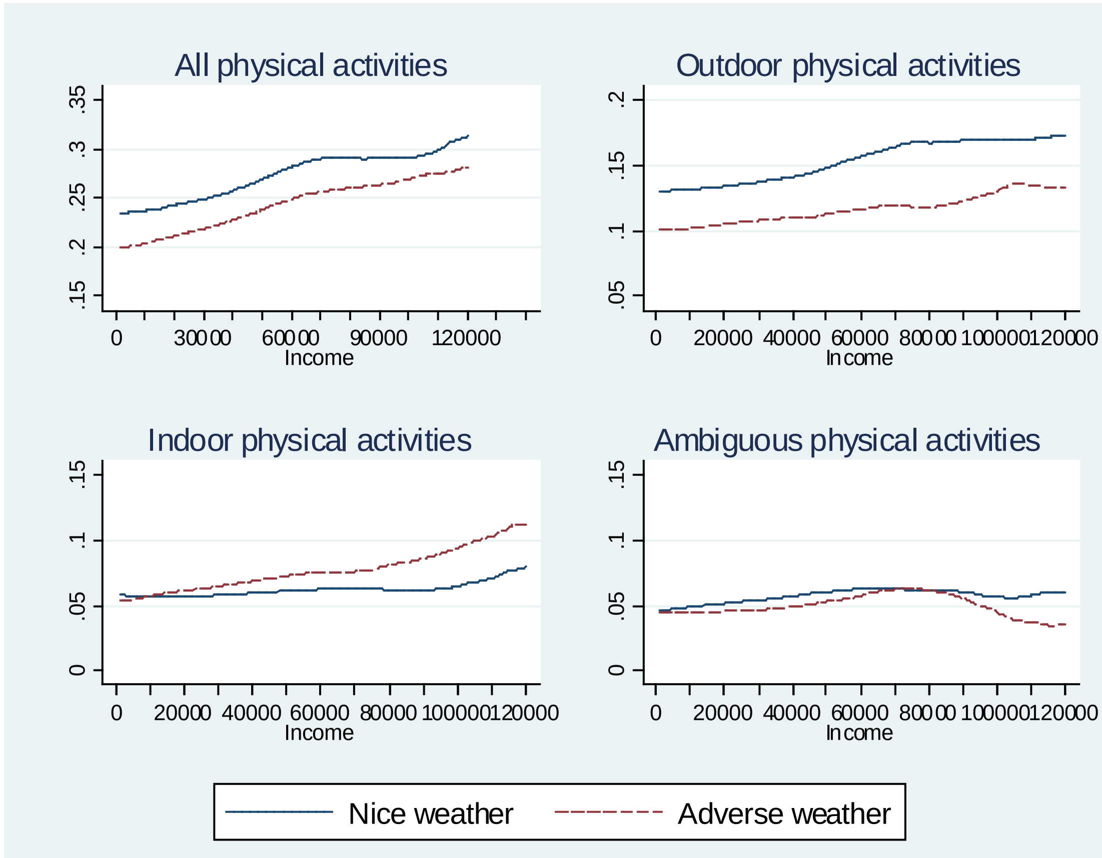
Source: Activity data from 1992, 1998 and 2005 General Social Survey (GSS), excluding respondents with activity limitations due to mental and/or physical health reasons. Weather data from Canadian National Climate Data and Information Archive (NCDIA).

Figure 4.1: Nonparametric analysis of probability of participating in leisure time physical activities, by month



Notes: Vertical axis plots the predicted probability of participating in leisure time physical activities from using kernel density nonparametric regression analysis. High and low income are defined as above the 75$^{\text{th}}$ and below the 25$^{\text{th}}$ percentiles in the sample, respectively.

Figure 4.2: Nonparametric analysis of probability of participating in leisure time physical activities, by income



Notes: Vertical axis plots the predicted probability of participating in leisure time physical activities from using kernel density nonparametric regression analysis. Nice weather is defined as existence of 0 hour of adverse weather between 6am to 11pm, and adverse weather is defined as existence of more than or equal to 1 hour of adverse weather between 6am and 11pm.

Conclusion

The overriding theme of my dissertation is the use of short-term weather fluctuations to study people's time allocation decisions, in the hope of providing valuable insights for understanding certain social phenomena behind these decisions. The theoretical foundation of weather-induced changes in time allocations is rooted in standard rational choice theory in economics, which, in this context, suggests that people allocate their limited time, on the margin, to the activities from which they derive the highest marginal utility. Weather significantly affects people's time allocation decisions through its varied impacts on utility associated with different activities, with outdoor recreational activities being the most sensitive type of activities. Time reallocation, therefore, occurs when people substitute away from low- towards high-utility activities.

Two particular types of substitutions induced by weather are examined. The first three chapters are devoted to understanding the substitution away from work to leisure when the quality of outside weather improves. Chapter 1 develops a theoretical model to formally illustrate how weather improvements induce sickness absenteeism. Relying on the assumption that the marginal utility of indoor leisure depends on one's sickness level while the utility of outdoor leisure depends on the interaction between one's health and outside weather, this model provides a setting in which weather-induced sickness absenteeism is unambiguously malfeasant in nature. Moreover, a key proposition of the model is that weather-absence correlation is larger among employees who are facing lower existing shirking incentives. Given the vagueness of the distinction between legitimate and illegitimate absenteeism in the existing literature, the model in Chapter 1 not only clearly distinguishes malfeasant from legitimate employee sickness absenteeism in theory, but also suggests an empirical strategy to identify a shirking component in overall reported sickness absenteeism, with which employee shirking behaviour can be studied in depth.

In order to exploit exogenous weather fluctuations to identify shirking absenteeism, as suggested in Chapter 1, Chapter 2 constructs an index of weather quality, which quantifies the ideal weather conditions for engaging in high-utility outdoor recreational activities. The weather quality index is constructed taking into account the multifaceted nature of weather conditions and captures how various weather elements – temperature, humidity, precipitation, wind speed, and cloud cover – come together to affect the propensity of employees to engage in outdoor recreational activities. The effect of these weather variables on outdoor recreational activities is estimated by linking data on the daily activities of paid employees from the Canadian General Social to weather conditions the employees face at the precisely same point in time. The results shows that weather conditions have significant effects on the probability of employees engaging in outdoor activities. Because the theoretical foundation of weather-induced changes in labour supply comes from its large impacts on the marginal utility of outdoor leisure, the findings in this chapter give us some confidence of the meaningfulness of the index when we relate it to labour supply decisions.

Chapter 3 empirically tests for the occurrence of weather-induced substitutions between work and outdoor recreational activities and examines how this type of employee behaviour varies across workers facing different shirking incentives. Linking 12 years of microdata on the self-reported sickness absenteeism of paid employees from Canada's Labour Force Survey (LFS) with the weather quality index constructed in Chapter 2, we find clear evidence of a positive relationship in the non-winter months between the quality of outside weather conditions and reported short-term sickness absenteeism, suggesting that the substitution away from work to outdoor leisure exists. Moreover, consistent with a key proposition of the theoretical model in Chapter 1, the empirical relation between weather and sickness absenteeism tends to be larger when existing shirking incentives are low. There is, however, little evidence that firms are able to adjust shirking incentives through payments of efficiency wages.

Chapter 4 turns to examine another type of substitution induced by weather variations – the substitution between outdoor and indoor physical activities. The Chapter begins with a theoretical model of the decision to participate in physical activities, which assumes that when adverse weather shocks deter outdoor physical activities, indoor physical activities are the only viable option for individuals to stay physically active. However, because the indoor options are more costly, substituting from outdoor to indoor physical activities is easier for high-income individuals. This suggests an explanation for the stylized fact that rates of physical activity participation are low among lower socioeconomic groups. Linking time-use data from the Canadian General Social Survey with archival weather data, the results of the empirical analysis in this chapter provide evidence of a positive income effect enabling substitution from outdoor to indoor physical activities when outside weather is not conducive for participating in outdoor activities. By exploiting the role that income plays in maintaining physical activity levels when less costly outdoor options are limited, this chapter formally illustrates a credible causal link between people's income levels and their participation in leisure time physical activities and provides direct evidence of this link.

The findings in this dissertation have several important policy implications. In the case of substitution from work to leisure, the main finding predicted by the theoretical model that marginal improvements in weather quality have their largest impact where existing shirking levels are lowest, is also supported by the results of empirical analysis using real-world data. This finding is somewhat counterintuitive in the sense that we would in general expect to see people who are less likely to be fired for shirking to be more likely to take an illegitimate absence from work to exploit nice weather. However, it is the inconsistency with the popular perception about malfeasant absenteeism (or employee shirking behaviour in general) that makes the finding an important reminder for employers or policy makers to consider when designing policy changes that attempt to reduce workplace absenteeism. Within an organization, different job positions allow different levels of time flexibility, which may already lead to different absence behaviour among employees. For example, in a corporation, the sale representatives are more likely to have more time flexibility than the workers working on product assembly lines. Our finding suggests that in order to make policies most effective in reducing workplace absence, employers need to recognize employees in what types of job positions or which departments are currently at lower absence level because those employees' malfeasant absence behaviour may be most responsive to changes in shirking incentives.

In the case of substitution between outdoor and indoor physical activities when weather conditions are not conducive for outdoor activities, the main finding is not only consistent with the existing literature, in which a positive correlation is consistently found between income and participation in physical activities, but more importantly, the main finding here illustrates that one of the mechanisms through which higher income leads to more physical activities. Specifically, higher income provides easier access and more options for participating in physical activities. Since one of the causal links between income and physical activities is identified, policies that try to promote physical activities, especially among low-income people, could be designed to directly provide the conditions as suggested by the mechanism of the income effect. This provides an alternative to offering assistance to improve income levels through various channels. Differences in dieting habits have been commonly used to explain the health inequality across income groups. Given the benefits of physical activity to both physical and mental health, the finding in this dissertation also suggests that policies aimed to reduce health inequality should not ignore the variations in exercise behaviour as a contributor to inequality.

This dissertation takes an approach of using changes in time allocation induced by exogenous weather variations to study human behaviour. The premise for using this approach is that weather can cause substitution between activities, and the findings provide evidence that people's time allocation decisions do respond to weather changes, although the magnitude may be small. The fact that weather has an impact on time allocation and the exogenous nature of weather variations together raise the hope of identifying causal factors behind certain social behaviour.

References

Ajzen, I. (1991), "The Theory of Planned Behavior," *Organizational Behavior and Human Decision Processes* 50: 179-211.

Ajzen, I. and M. Fishbein (1980), "Understanding attitudes and predicting social behavior," Englewood Cliffs, NJ: Prentice-Hall.

Allen, Steven (1983), "How much does absenteeism cost," *Journal of Human Resources* 18(3): 379-393.

Angrist, Joshua D., Kathryn Graddy and Guido W. Imbens (2000), "The interpretation of instrumental variables estimators in simultaneous equations models," *Review of Economic Studies* 67(3): 499-527.

Arai, Mahmood and Peter Skogman Thoursie (2005), "Incentives and selection in cyclical absenteeism," *Labour Economics* 12:269-280.

Askildsen, Jan Erik, Espen Bratberg and Øivind Anti Nilsen (2005), "Unemployment, labor force composition and sickness absence: a panel data study," *Health Economics* 14: 1087-1101.

Audas, Rick and John Goddard (2001), "Absenteeism, seasonality, and the business cycle," *Journal of Economics and Business* 53: 405-419.

Azevedo, Mario Renato, Cora Luiza Pavin Arajo, Felipe Fossati Reichert, Fernando Vinholes Siqueira, Marcelo Cozzensa da Silva, and Pedro Curi Hallal (2006), "Gender differences in leisure-time physical activity," *International Journal of Public Health* 52(1): 8-15.

Ball, Kylie, Jo Salmon, Billie Giles-Corti, and David Crawford (2006), "How Can Socio-economic Differences in Physical Activity among Women be Explained? A Qualitative Study," *Women Health* 43(1): 93-113.

Barmby, Tim, Chris Orme, and John Treble (1995), "Work absence histories: a panel data study," *Labour Economics* 2, 53-65.

Barmby, Tim, John Session, and Jphn Treble (1994), "Absenteeism, efficiency wages and shirking," *Scandinavian Journal of Economics* 96(4), 561-566.

Bauman, Adrian, Guansheng Ma, Frances Cuevas, Zainal Omar, Temo Waqanivalu, Philayrath Phongsavan, Kieren Keke, Anjana Bhushan, for the Equity and Non-communicable Disease Risk Factors Project Collaborative Group (2011), "Cross-national Comparisons of Socioeconomic Differences in the Prevalence of Leisure-time and Occupational Physical Activity, and Active Commuting in Six Asia-Pacific Countries," *Journal of Epidemiology and Community Health* 65: 35-43.

Becker G.S. (1965), "A Theory of the Allocation of Time," *The Economic Journal* 75: 493-513.

Booth, M.L., A. Bauman, N. Owen, and C.J. Gore (1997), "Physical Activity Preferences, Preferred Sources of Assistance, and Perceived Barrier to Increased Activity among Physically Inactive Australian," *Preventive Medicine* 26(1): 131-137.

Bradley, Steve, Colin Green, and Gareth Leeves (2007), "Worker absence and shirking: evidence from matched teacher-school data," *Labour Economics* 14: 319-334.

Brown, Heather, and Jennifer Roberts (2011), "Exercising Choice: The Economic Determinants of Physical Activity Behaviour of an Employed Population," *Social Science and Medicine* 73: 383-390.

Capelli, Peter and Keith Chauvin (1991), "An Interplant Test of the Efficiency Wage Hypothesis," *Quarterly Journal of Economics* 106(3): 769-787.

Carlin, Paul (1989), "Why the incidence of shirking varies across employees," *Journal of Behaviour Economics* 18(2), 61-73.

Cawley, John (2004), "An Economic Framework for Understanding Physical Activity and Eating Behavors," *American Journal of Preventive Medicine* 27: 117-125.

Cerin, Ester and Eva Leslie (2008), "How Socio-economic Status Contribute to Participation in Leisure-time Physical Activity," *Social Science and Medicine* 66: 2596-2609.

Cerveny, R.S. and R.C. Balling Jr. (1998), "Weekly cycles of air pollutants, precipitation and topical cyclones in the coastal NW Atlantic region," *Naure* 394: 561-563.

Charles, Kerwin Kofi and Philip Decicca (2008), "Local labour market fluctuations and health: is there a connection and for whom?" *Journal of Health Economics* 27: 1532-1550.

Chatterji, Monojit and Colin Tilley (2002), "Sickness, absenteeism, presenteeism, and sick pay," *Oxford Economic Papers* 54, 669-687.

Chen, Paul and Per-Anders Edin (2002), "Efficiency Wages and Industry Wage Differentials: A Comparison Across Methods of Pay," *Review of Economics and Statistics* 84(4): 617-631.

Chinn, David J., Martin White, Jane Harland, Christopher Drinkwater, and Simon Raybould (1999), "Barriers to physical activity and socioeconomic position: implications for health promotion," *Journal of Epidemiol Community Health* 52: 191-192.

Connolly, Marie (2008), "Here comes the rain again: weather and the intertemporal substitution of leisure," *Journal of Labor Economics* 26(1): 73-100.

Craig, C.L., S.J. Russell, C. Cameron, and A. Bauman (2004), "Twenty-year Trends in Physical Activity among Canadian Adults," *Canadian Journal of Public Health* 95(1): 59-63.

De Freitas, C.R., Daniel Scott, and Geoff McBoyle (2008), "A second generation climate index for tourism (CIT): specification and verification," *International Journal of Biometeorology* 52(5): 399-407.

Dionne, Georges and Benoit Dostie (2007), "New evidence on the determinants of absenteeism using linked employer-employee data," *Industrial and Labor Relations Review* 61(1): 108-119.

Droomers, M., C.T.M. Schrijvers, and J.P. Mackenhach (2001), "Educational Level and Decrease in Leisure Time Physical Activity: Predictors from the Longitudinal GLOBE Study," *Epidemiology and Community Health* 55: 562-568.

Dunn, L.F. and Stuart Youngblood (1986), "Absenteeism as a mechanism for approaching optimal labour market equilibrium: an empirical study," *Review of Economics and Statistics* 68(4), 668-674.

Eisengerg, Daniel and Edward Okeke (2009, "Too cold for a jog? weather, exercise, and socioeconomic status," *The B.E. Journal of Economic Analysis and Policy* 9(1): 1-30.

Estabrooks, P.A., R.E. Lee, and N.C. Gyurcsik (2003), "Resources for Physical Activity Participation: Does Availability and Accessibility Differ by Neighborhood Socioeconomic Status?" *Annals of Behavioral Medicine* 25: 100-104.

Frick, Bernd and Miguel Malo (2008), "Labor market institutions and individual absenteeism in the European union: the relative importance of sickness benefit systems and employment protection legislation, *Industrial Relations* 47(4): 505-529.

Gilmour, Heather (2007), "Physically Active Canadians," *Health Reports* vol.18, no.3, August, Statistics Canada Catalogue no. 82-003: 45.

Goetzel, RZ and RJ Ozminkowski (2008), "The health and cost benefits of work site health-promotion programs," *Annual Review of Public Health* 29: 303-323.

Harrison, D.A. and J.J. Martocchio (1998), "Time of absenteeism: a 20-year review of origins, offshoots, and outcomes," *Journal of Management* 24(3): 305-350.

Howard, E. and M.S. Hoffman (1984), "A multidimensional approach to the relationship between mood and weather," *British Journal of Psychologu* 75(1): 12-23.

Huang, Tzu-Ling, Arne Hallam, Peter F. Orazem, and Elizabeth M. Paterno (1998), "Empirical Tests of Efficiency Wage Models," *Economica* 65(257): 125-143.

Humbert, M. Louise, Karen E. Chad, Kevin S. Spink, Nazeem Muhajarine, Kristal D. Anderson, Mark W. Bruner, Tammy M. Girolami, Patrick Odnokon, and Catherine R. Gryba (2006), "Factors That Influence Physical Activity Participation Among High- and Low-SES Youth," *Qualitative Health Research* 16: 467-483.

Humphreys, Brad R. and Jane Ruseski (2010), "The Economic Choice of Participation and Time Spent in Physical Activity and Sport in Canada," *University of Alberta Working Paper* 2010-14.

Ichino, Andrea and Enrico Moretti (2006), "Biological gender differences, absenteeism and the earnings gap," NBER Working Paper Series, no. 12369.

Ichino, Andrea and Regina Riphahn (2005), "The effect of employment protection on worker effort: absenteeism during and after probation," *Journal of the European Economic Association* 3(1): 120-143.

Jacobsen, Ben and Wessel Marquering (2008), "Is it the weather?" *Journal of Banking and Finance* 32: 526-540.

Kanstra, Mark, Lisa Kramer and Maurice Levi (2003), "Winter blues: a SAD stock market cycle," *American Economic Review* 93(1): 324-343.

Katzmarzyk, Peter T. and Ian Janssen (2004), "The Economic Costs Associated with Physical Inactivity and Obesity in Canada: An Update," *Canadian Journal of Applied Physiology* 29: 90-115.

Kavanagh, A.M., J.L. Goller, T. King, D. Jolley, D. Crawford, and G. Turell (2005), "Urban Area Disadvantage and Physical Activity: A Multilevel Study in Melbourne, Australia," *Journal of Epidemiology and Community Health* 59: 934-940.

Kliger, Doron and Ori Levy (2003), "Mood-induced variation in risk preferences," *Journal of Economic Behaviour and Organization* 52: 573-584.

Lawler, Edward E. III. (1973), *"Motivation in Work Organizations,"* Monterey: Brooks/Cole.

Leigh, J.P. (1985), "The effects of unemployment and the business cycle on absenteeism," *Journal of Economics and Business* 37: 159-170.

Leonard, Jonathan S. (1987), "Carrots and Sticks: Pay, Supervision, and Turnover," *Journal of Labor Economics* 5(4): S136-S152.

Levy, Ori and Itai Galili (2008), "Stock purchases and the weather: individual differences," *Journal of Economic Behavior and Organization* 67: 755-767.

Lindstrom, M., B.S. Hanson, P.O. Ostergren (2001), "Socioeconomic Differences in Leisure-time Physical Activity: The Role of Social Participation and Social Capital in Shaping Health Related Behaviour," *Social Science and Medicine* 52: 441-451.

MacLeod, W. Bentley and James M. Malconsom (1998), “Motivation and Markets,” *American Economic Association* 88(3): 388-411.

Maddison, D. (2001) “In search of warmer climates? The impact of climate change on flows of British tourists,” *Clim Change* 49: 193-208.

Markham, Steven E. and Ina S. Markham (2005), “Biometeorolgical effects on worker absenteeism,” *International Journal of Biometeorolgy* 49: 317-324.

McNeil, L.H., M.W. Kreuter, and S.V. Subramanian (2006), “Social Environment and Physical Activity: A Review of Concepts and Evidence,” *Social Science and Medicine* 63: 1011-1022.

Merrill R.M., E.C. Shields, G.L. White, and D. Druce (2005), “Climate conditions and physical activity in the United States,” *American Journal of Health Behavior* 29: 371-381.

Mieczkowski, Z (1985) “The tourism climatic index: a method of evaluating world climates for tourism,” *Canadian Geographer* 29(3): 220-233.

Morgan, R., E. Gatell, R. Junyent, A. Micallef, E. Özhan, and A.T. Williams (2000) “An improved user-based beach climate index,” *Journal of Coastal Conversation* 6: 41-50.

Moschini, Giancarlo and David A. Hennessy (2001), “Uncertainty, risk aversion, and risk management for agricultural producers,” in *Handbook of Agricultural Economics*, Bruce L. Gardner and Gordon C. Rausser, eds., Amsterdam, North-Holland, Vol. 1, Part 1, 88-153.

Mueser, Roland E. (1953), “The weather and other factors influencing employee punctuality,” *Journal of Applied Psychology* 37(5): 329-337.

Murphy, Kevin M. and Robert H. Topel (1990), “Efficiency Wages Reconsidered: Theory and Evidence,” in Y. Weiss and G. Fishelson, *Advances in The Theory and Measurement of Unemployment*, New York: St. Martin’s Press: 204-240.

Neal, Derek (1993), “Supervision and Wages Across Industries,” *Review of Economics and Statistics* 75(3): 409-17.

Nicholson, Sean, Mark Pauly, Daniel Polsky, Catherine Basse, Gary Billotti, Ronald Ozminkowski, Marc Berger, and Claire Sharda (2005), “How to present the business case for healthcare quality to employers,” *Applied Health Economics and Health Policy* 4(4), 209-218.

Paringer, Lynn (1983), “Women and absenteeism: health or economics?” *American Economic Review* 73(2): 123-27.

Parks, S.E., R.A. Housemann, and R.C. Brownson (2003), “Differential Correlates of Physical Activity in Urban and Rural Adults of Various Socioeconomic Backgrounds in the United States,” *Journal of Epidemiology and Community Health* 57: 29-35.

Paxson, Christina H. (1992), "Using weather variability to estimate the response of savings to transitory income in Thailand," *American Economic Review* 82(1): 15-33.

Pocock, S.J. (1972), "Relationship between sickness absence and meteorological factors," *British Journal of Preventive and Social Medicine* 26: 283-45.

Powell, L.M., S. Slater, and F.J. Chaloupka (2004), "The Relationship between Community Physical Activity Settings and Race, Ethnicity and Socio-economic Status," *Evidence-based Preventive Medicine* 1: 135-144.

Roll, Richard (1984), "Orange juice and weather," *American Economic Review* 74(5): 861-880.

Ruhm, Christopher J. (2000), "Are recessions good for your health?" *Quarterly Journal of Economics* 115(2): 617-50.

Sallis, James F., Adrian Bauman, and Michael Pratt (1998), "Environmental and policy interventions to promote physical activity," *American Journal of Preventive Medicine* 15(4): 379-397.

Samuelson, P. (1938), "A Note on the Pure Theory of Consumers' Behaviour," *Economica* 5: 61-71.

Sari, Nazmi (2009), "Physical Inactivity and Its Impact on Healthcare Utilization," *Health Economics* 18(8): 885-901.

Saunders, Edward (1993), "Stock prices and Wall Street weather," *American Economic Review* 83(5): 1337-1345.

Scott, Daniel and Geoff McBoyle (2001), "Using a "tourism climate index" to examine the implications of climate change for climate as a natural resource for tourism," *Proceedings of the First International Workshop on Climate, Tourism and Recreation. International Society of Biometeorology, Commission on Climate, Tourism and Recreation, Halkidi, Greece* 69C98.

Shapiro, Carl and Joseph E. Stiglitz (1984), "Equilibrium unemployment as a worker discipline device," *American Economic Review* 74: 433-444.

Starr-McCluer, Martha (2000), "The effects of weather on retial sales," FEDS Working Paper No. 2000-08, Federal Reserve Board of Governors, Washing D.C.

Steers, R.M. and S.R. Rhodes (1978), "Major influences on employee attendance: a process model," *Journal of Applied Psychology* 63: 391-407.

Sternfeld, B., B.E. Ainsworth, and C.P. Quesenberry Jr. (1999), "Physical Activity Patterns in a Diverse Population of Women," *Preventive Medicine* 28(3): 313-323.

Vistnes, Jessica Primoff (1997), "Gender differences in days lost from work due to illness," *Industrial and Labor Relations Review* 50(2): 304-523.

Wadhwani, Sushil B. and Martin Wall (1991), “A Direct Test of the Efficiency Wage Model Using UK Micro-Data,” *Oxford Economic Papers* 43(4): 529-548.

Warburton, Darren E.R., Crystal W. Nicol and Shannon S. D. Bredin (2006), “Health Benefits of Physical Activity: the Evidence,” *Canadian Medical Association Journal* 176: 801-808.

Wilson, D.K., K.A. Kirkland, B.E. Ainsworth, and C.L. Addy (1992), “Socioeconomic Status and Perceptions of Access and Safety for Physical Activity,” *Annals of Behavioral Medicine* 28: 20-28.

Yen, I.H. and G.A. Kaplan (1998), “Poverty Area Residence and Changes in Physical Activity Level: Evidence from the Alameda County Study,” *American Journal of Public Health*: 88(11): 1709-1212.

Zivin, Joshua G. and Matthew J. Neidell (2010), “Temperature and the allocation of time: implications for climate change,” NBER Working paper No. 15717.

APPENDICES

Appendix A

The Appendix for Chapter 1

Continuation Values

At any one time period, the utilities for a worker being "not-absent from work", "absent from work", and unemployed are:

$$u = \begin{cases} u^{na} = (1-\delta)\, w + \delta\, (T-h), & \text{if not absent in that period} \\ u^{a} = (1-\delta)\, s + \delta\, T, & \text{if absent and not dismissed in that period} \\ u^{u} = (1-\delta)\, b + \delta\, T, & \text{if absent and dismissed in that period} \end{cases}$$

Where δ is marginal utility of leisure, which equals to θ for indoor leisure and $(1-\theta)\lambda$ for outdoor leisure. Under weather λ, when a worker's sickness level above θ^i and below θ^o, the worker will prefer being absent from work, enjoying indoor leisure and outdoor leisure, respectively. θ^z defines the threshold above which the sickness absence is legitimate and the worker is not facing a risk of getting dismissed for taking paid sick leave. Given the level of sickness, θ, is uniformly distributed over the interval [0,1], the expected utilities of a worker being employed and unemployed in any period, respectively, are:

$$\begin{aligned} E(U^e) &= \int_{\theta^z}^{1} [(1-\theta)\, s + \theta\, T]\, d\theta + \int_{\theta^i}^{\theta z} \{\alpha\ [(1-\theta)\, b + \theta\, T] + (1-\alpha)\ [(1-\theta)\, s + \theta\, T]\} d\theta \\ &= \int_{\frac{\lambda}{1+\lambda}}^{\theta^i} [(1-\theta)\, w + \theta\, (T-h)] d\theta + \int_{\theta^o}^{\frac{\lambda}{1+\lambda}} [(1-\theta\lambda)\, w + \theta\lambda\, (T-h)] d\theta \\ &= + \int_{0}^{\theta^o} \{\alpha\ [(1-\theta\lambda)\, b + \theta\lambda\, T] + (1-\alpha)\ [(1-\theta\lambda)\, s + \theta\lambda\, T]\} d\theta \end{aligned}$$

and,

$$\begin{aligned}
E(U^u) &= \int_{\theta^z}^{1} [(1-\theta)\, b + \theta\, T]\, d\theta + \int_{\theta^i}^{\theta z} [(1-\theta)\, b + \theta\, T] d\theta \\
&= \int_{\frac{\lambda}{1+\lambda}}^{\theta^i} [(1-\theta)\, b + \theta\, T]\, d\theta + \int_{\theta^o}^{\frac{\lambda}{1+\lambda}} [(1-\theta\lambda)\, b + \theta\lambda\, T] d\theta \\
&= + \int_{0}^{\theta^o} [(1-\theta\lambda)\, b + \theta\lambda\, T]\, d\theta.
\end{aligned}$$

Since $s > b$, It is obvious that $[(1-\theta)\, s + \theta\, T] > [(1-\theta)\, b + \theta\, T]$, $\{\alpha\, [(1-\theta)\, b + \theta\, T] + (1-\alpha)\, [(1-\theta)\, s + \theta\, T]\} > [(1-\theta)\, b + \theta\, T]$, and $\{\alpha\, [(1-\theta\lambda)\, b + \theta\lambda\, T] + (1-\alpha)\, [(1-\theta\lambda)\, s + \theta\lambda\, T]\} > [(1-\theta\lambda)\, b + \theta\lambda\, T]$; also, because when a worker is employed, they show up at work when utility received from work is greater than from leisure, and this implies $[(1-\theta)\, b + \theta\, T] > [(1-\theta)\, b + \theta\, T]$ and $[(1-\theta\lambda)\, b + \theta\lambda\, T] > [(1-\theta\lambda)\, b + \theta\lambda\, T]$. Hence, it is always the case that $[E(U^e) - E(U^u)] > 0$.

The continuation values for a worker at the beginning of period 2 with and without a contract, are respectively,

$$\begin{aligned}
V(E) &= E(U^e) + \rho\, \left[1 - \alpha\, (\theta^o + \theta^z - \theta^i)\right] V_3(E) + \rho\, \alpha\, (\theta^o + \theta^z - \theta^i) V_3(U) \\
V(U) &= E(U^u) + \, a\, V_3(E) + \rho\, (1-a)\, V_3(U)
\end{aligned}$$

Therefore,

$$V(E) - V(U) = E(U^e) - E(U^u) + \rho\, \left[1 - \alpha\, (\theta^o + \theta^z - \theta^i)\right] [V_3(E) - V_3(U)]$$

where $V_3(E)$ and $V_3(U)$ are the continuation values at the beginning of period 3 being employed and unemployed, respectively. Following this derivation process, over an infinite time horizon, the difference in the continuation values between being employed and unemployed at the beginning of period 2 is:

$$\begin{aligned}
V(E) - V(U) &= \sum_{t=1}^{\infty} \rho\, \left[1 - \alpha\, (\theta^o + \theta^z - \theta^i) - a\right]^{t-1} [E(U^e) - E(U^u)] \\
&= \frac{1}{1 - \rho\, [1 - \alpha\, (\theta^o + \theta^z - \theta^i) - a]} [E(U^e) - E(U^u)]
\end{aligned}$$

which is necessarily positive.

Appendix B

The Appendix for Chapter 3

Identification of Outdoor Workers

Four-digit industries: Oilseed and grain farming; Vegetable and melon farming; Fruit and tree nut farming; Greenhouse, nursery and floriculture production; Other crop farming; Cattle ranching and farming; Hog and pig farming; Poultry and egg production; Sheep and goat farming; Animal aquaculture; Other animal production; Timber tract operations; Forest nurseries and gathering of forest products; logging; Fishing; Hunting and trapping; Support activities for crop production; Support activities for animal production; Support activities for forestry; Oil and gas extraction; Coal mining; Metal ore mining; Non-metallic mineral mining and quarrying; Support activities for mining and oil and gas extraction; Electric power generation, transmission and distribution; Natural gas distribution; Water, sewage and other systems; Residential building construction; Non-residential building construction; Utility system construction; Land subdivision; Highway, street and bridge construction; Other heavy and civil engineering construction; Foundation, structure, and building exterior contractors; Building equipment contractors; Building finishing contractors; other specialty trade contractors; Scheduled air transportation; Non-scheduled air transportation; Rail transportation; Deep sea, coastal and great lakes water transportation; Inland water transportation; General freight trucking; Specialized freight trucking; Urban transit systems; Interurban and rural bus transportation; Taxi and limousine service; School and employee bus transportation; Charter bus industry; Other transit and ground passenger transportation; Pipeline transportation of crude oil; Pipeline transportation of natural gas; other pipeline transportation; Scenic and sightseeing transportation, land; Scenic and sightseeing transportation, water; Scenic and sightseeing transportation, other; Support activities for air transportation; Support activities for rail transportation; Support activities for water transportation; Support activities for road transportation; Freight transportation arrangement; Other support activities for transportation; Postal service; Couriers; Local messengers and local delivery; Warehousing and storage; Services to building and dwellings; Waste collection; Waste treatment and disposal; Remediation and other waste management services; Spectator sports; Heritage institutions; Amusement parks and arcades; Other amusement and recreation industries; Recreational vehicle parks and recreational camps.

Four-digit occupations: Mail, postal and related clerks; Letter carriers; Couriers, messengers and door-to-door distributors; Land surveyors; Farmers and farm managers; Agricultural and related service contractors and managers; Farm supervisors and specialized livestock workers; Nursery and greenhouse operators and managers; Landscaping and grounds maintenance contractors and managers; Supervisors, landscape and horticulture; Aquaculture operators and managers; General farm workers; Nursery and greenhouse workers; Supervisors, logging and forestry; Supervisors, mining and quarrying; Supervisors, oil and gas drilling and service; Underground production and development miners; Oil and gas well drillers, services, testers and related workers; Underground mine service and support workers; Oil and gas well drilling workers and service operators; Logging machinery operators; Chainsaw and skidder operators; Silviculture and forestry workers; Fishing Masters and Officers; Fishing vessel skippers and Fishermen/women; Fishing vessel deckhands; Trappers and hunters; Harvesting labourers; Landscaping and grounds maintenance labourers; Aquaculture and marine harvest labourers; Mine labourers; Oil and gas drilling, servicing and related labourers; Logging and Forestry labourers; Tour and travel guides; Outdoor sport and recreational guides; Heavy equipment operators; Public works maintenance equipment operators; Crane operators; Drillers and blasters; Water well drillers; Truck drivers; Bus drivers and subway and other transit operators; Taxi and limousine drivers and chauffeurs; Delivery and courier service drivers; Railway and yard locomotive engineers; Railway conductors and brakemen/women; Railway yard workers; Railway track maintenance workers; Deck crew, water transport; Engine room crew, water transport; Lock and cable ferry operators and related occupations; Boat operators; Air transport ramp attendants.

Appendix C

The Appendix for Chapter 4

1. Optimal decision of time allocation and consumption when $\lambda > \lambda_0$

When $\lambda > \lambda_0$, outdoor LTPA is preferred to indoor LTPA, and an individual solves the following problem to maximize their his or her utility:

$$\max_{t_o,t_z,C} U\{t_o,t_z,C\} = [I(\lambda)kt_o + (1-k)t_z^{\alpha}]^{\eta}C^{1-\eta} \tag{C.0.1}$$

subject to

$$p_e t_o + p_z t_z + C = Y_0 + wT \tag{C.0.2}$$

where t_o and t_z are the time spent on outdoor LTPA and other non-market activities, and p_e and p_z are the prices associated with outdoor LTPA and the non-market activities respectively. C is the consumption on commodities. Y_0 is the non-labour income income and w is the wage rate. The solution of the maximization problem can be obtained by solving the following Lagrangian problem:

$$L = [kt_o + (1-k)t_z^{\alpha}]^{\eta}C^{1-\eta} - \gamma(p_e t_o + p_z t_z + C - Y_0 - wT) \tag{C.0.3}$$

where γ is the Lagrangian multiplier. Because of the quailinear feature of the utility function, when the following first order conditions of the Lagrangian are satisfied, the optimal solution for t_o is zero:

$$F.O.C. = \begin{cases} \frac{\partial L}{\partial t_o} = \eta[(1-k)t_z^{\alpha}]^{\eta-1}C^{1-\eta}k - \gamma p_e < 0 \\ \frac{\partial L}{\partial t_z} = \eta[(1-k)t_z^{\alpha}]^{\eta-1}C^{1-\eta}(1-k)\alpha t_z^{\alpha-1} - \gamma p_z = 0 \\ \frac{\partial L}{\partial C} = [(1-k)t_z^{\alpha}]^{\eta}(1-\eta)C^{-\eta} - \gamma = 0 \\ \frac{\partial L}{\partial \gamma} = Y_0 + wT - p_z t_z - C = 0 \end{cases}$$

Therefore, in this situation, the solution for the maximization problem is:

$$(t_o^*, t_z^*, C) = \begin{cases} t_o^* = 0 \\ t_z^* = \frac{\alpha\eta(Y_0+wT)}{(1-\eta+\alpha\eta)p_z} \\ C^* = \frac{(1-\eta)(Y_0+wT)}{1-\eta+\alpha\eta} \end{cases}$$

If the optimal value of t_o is greater than 0, the following first order conditions must hold:

$$F.O.C. = \begin{cases} \frac{\partial L}{\partial t_o} = \eta[kt_o + (1-k)t_z^{\alpha}]^{\eta-1}C^{1-\eta}k - \gamma p_e = 0 \\ \frac{\partial L}{\partial t_z} = \eta[kt_o + (1-k)t_z^{\alpha}]^{\eta-1}C^{1-\eta}(1-k)\alpha t_z^{\alpha-1} - \gamma p_z = 0 \\ \frac{\partial L}{\partial C} = [kt_o + (1-k)t_z^{\alpha}]^{\eta}(1-\beta)C^{-\eta} - \gamma = 0 \\ \frac{\partial L}{\partial \gamma} = Y_0 + wT - p_e t_o - p_z t_z - C = 0 \end{cases}$$

Solving for the above F.O.C., the optimal values of t_o, t_z, and C are:

$$(t_o^*, t_z^*, C) = \begin{cases} t_o^* = \frac{\eta(Y_0+wT)}{p_e} - \frac{(1-k)^{\frac{1}{1-\alpha}}[\eta\alpha^{\frac{1}{1-\alpha}} p_e^{\frac{1}{1-\alpha}} + (1-\eta)\alpha^{\frac{\alpha}{1-\alpha}} p_e^{\frac{\alpha}{1-\alpha}}]}{k^{\frac{1}{1-\alpha}} p_z^{\frac{\alpha}{1-\alpha}}} \\ t_z^* = [\frac{(1-k)\alpha p_e}{kp_z}]^{\frac{1}{1-\alpha}} \\ C^* = \frac{(1-\eta)p_e[kt_o^* + (1-k)t_z^{*\alpha}]}{k\eta} \end{cases}$$

2. Optimal decision of time allocation and consumption when $\lambda < \lambda_0$

When $\lambda < \lambda_0$, indoor LTPA is preferred to outdoor LTPA, and an individual solves the following problem to maximize their his or her utility:

$$\max_{t_i, t_z, C} U\{t_i, t_z, C\} = [I(\lambda)kt_i + (1-k)t_z^{\alpha}]^{\eta}C^{1-\eta} \tag{C.0.4}$$

subject to

$$F_i + p_e t_i + p_z t_z + C = Y_0 + wT \tag{C.0.5}$$

where t_i is the time spent on indoor LTPA, and F_i is the access fee associated with indoor LTPA. The solution of the maximization problem can be obtained by solving the following Lagrangian problem:

$$L = [kt_i + (1-k)t_z^{\alpha}]^{\eta}C^{1-\eta} - \gamma(F_i + p_e t_i + p_z t_z + C - Y_0 - wT) \tag{C.0.6}$$

where γ is the Lagrangian multiplier. Because of the quailinear feature of the utility function, when the following first order conditions of the Lagrangian are satisfied, the optimal solution for t_i is zero:

$$F.O.C. = \begin{cases} \frac{\partial L}{\partial t_i} = \eta[(1-k)t_z^{\alpha}]^{\eta-1}C^{1-\eta}k - \gamma p_e < 0 \\ \frac{\partial L}{\partial t_z} = \eta[(1-k)t_z^{\alpha}]^{\eta-1}C^{1-\eta}(1-k)\alpha t_z^{\alpha-1} - \gamma p_z = 0 \\ \frac{\partial L}{\partial C} = [(1-k)t_z^{\alpha}]^{\eta}(1-\eta)C^{-\eta} - \gamma = 0 \\ \frac{\partial L}{\partial \gamma} = Y_0 + wT - F_i - p_z t_z - C = 0 \end{cases}$$

Therefore, in this situation, the solution for the maximization problem is:

$$(t_i^*, t_z^*, C) = \begin{cases} t_i^* = 0 \\ t_z^* = \frac{\alpha\eta(Y_0+wT-F_i)}{(1-\eta+\alpha\eta)p_z} \\ C^* = \frac{(1-\eta)(Y_0+wT-F_i)}{1-\eta+\alpha\eta} \end{cases}$$

If the optimal value of t_i is greater than 0, the following first order conditions must hold:

$$F.O.C. = \begin{cases} \frac{\partial L}{\partial t_i} = \eta[kt_i + (1-k)t_z^{\alpha}]^{\eta-1}C^{1-\eta}k - \gamma p_e = 0 \\ \frac{\partial L}{\partial t_z} = \eta[kt_i + (1-k)t_z^{\alpha}]^{\eta-1}C^{1-\eta}(1-k)\alpha t_z^{\alpha-1} - \gamma p_z = 0 \\ \frac{\partial L}{\partial C} = [kt_i + (1-k)t_z^{\alpha}]^{\eta}(1-\eta)C^{-\eta} - \gamma = 0 \\ \frac{\partial L}{\partial \gamma} = Y_0 + wT - F_i - p_e t_i - p_z t_z - C = 0 \end{cases}$$

Solving for the above F.O.C., the optimal values of t_i, t_z, and C are:

$$(t_i^*, t_z^*, C) = \begin{cases} t_i^* = \frac{\eta(Y_0+wT-F_i)}{p_e} - \frac{(1-k)^{\frac{1}{1-\alpha}}[\eta\alpha^{\frac{1}{1-\alpha}} p_e^{\frac{1}{1-\alpha}} + (1-\eta)\alpha^{\frac{\alpha}{1-\alpha}} p_e^{\frac{\alpha}{1-\alpha}}]}{k^{\frac{1}{1-\alpha}} p_z^{\frac{\alpha}{1-\alpha}}} \\ t_z^* = [\frac{(1-k)\alpha p_e}{kp_z}]^{\frac{1}{1-\alpha}} \\ C^* = \frac{(1-\eta)p_e[kt_o^* + (1-k)t_z^{*\alpha}]}{k\eta} \end{cases}$$