INTRODUCTION

Improving the living standards of populations is a widespread societal objective. A cornerstone of living standards is the ability of individuals to earn wages and profits in order to purchase goods and services for consumption. In turn, wages and profits reflect the value of the goods and services produced in an economy and the productivity of the factor inputs used to produce them. Though living standards, income and productivity are distinct concepts, the three are very much related. The correlation between labour productivity and real wages both across countries and over time is quite high, indicating the importance of productivity growth rates for the improvement of a country’s living standards (Harris 1999).

Consequently, economists and historians have focused much attention on better understanding the determinants of productivity growth. There is increasing awareness that human capital is a key factor. Traditionally, human capital has been interpreted as education and skills. Recently, however, increasing attention has been given to health as a form of human capital.

Over the last few years a growing body of literature has developed on the macroeconomic and microeconomic relationship between health and productivity. This chapter reviews the theoretical underpinnings and empirical evidence of this relationship. In particular, it addresses the question: Would an improvement in the health status of working Canadians pay off in terms of higher aggregate productivity? The evidence presented comes from many countries, both developed and developing, and spans a period of over 200 years. The review focuses on implications for public policy and firm-level practices in developed countries, particularly Canada.

Figure 1 provides an organizing framework for this review. It lists a number of strategies that the public and private sectors have employed to promote the health of individuals and populations, as well as several measures of health to assess the effectiveness of these strategies. Some of the strategies are employed with the express intention of improving human capital and, in turn, productivity. Others have improving health as a specific goal. The more traditional strategies of sanitation, nutrition and education are...
public-sector interventions targeted at population health, but they nonetheless have implications for the health and productivity of the labour force. Another traditional public-sector intervention, occupational health and safety, is targeted at the workplace and is focused on the reduction of accidents and chemical exposures and the resultant work-related injuries, illnesses and disabilities. The strategies of health promotion and healthy-workplace promotion are newer, firm-level initiatives developed since the 1970s from the growing awareness that organizational-level interventions can be an effective means of promoting healthy lifestyles, reducing stress, improving employee wellness, and reducing sickness-related absence and health-care costs (Polanyi et al. 2000). Lastly, population health is a new strategy based on the acknowledgement that the determinants of health are multifactorial — biological, social and economic — and that health policy needs to take a broad, multisectoral approach (Frank 1995).

Figure 1 also lists several labour-productivity and standard-of-living measures that can be affected by improvements in the health of the labour force. At the individual level, health can directly increase general output (e.g., through enhanced physical energy and mental acuity), yearly output (e.g., through reduced sickness absence) and career output (e.g., through decreased morbidity or increased longevity, resulting in a longer career). At the aggregate level, these individual increases in output can translate into increases in labour productivity (i.e., output per hour worked, output per worker) and/or standard of living (i.e., GNP per capita) (e.g., by increasing the size of the active labour force relative to the population).

This chapter proceeds as follows: The next section, “Human Capital, Health and Productivity,” reviews the theory on the demand for health and its relationship to the accumulation of human capital. “Historical Trends and Current Macroeconomic Evidence” reviews the historical economic evidence concerning the relationship between health and productivity growth as well as current macroeconomic empirical work on measuring this relationship. “Occupational Health and Safety and the Cost of Work Disability” reviews the changing nature of work and the implications for traditional approaches to occupational health and safety regulation — its ability to influence firm and worker behaviours and, through these, health and productivity. “Health, Sickness Absence and Firm-Level Practices” reviews the microeconomic evidence concerning the relationship between health and various productivity markers, with a focus on sickness absence.
absence. Lastly, the key implications for policy and future research are summarized.

HUMAN CAPITAL, HEALTH AND PRODUCTIVITY

Grossman's (1972, 2000) model for health demand provides insights into the relationship among health, human capital and consumption at the individual level, as well as a framework for modelling human capital accumulation and its relationship to productivity at the micro and macro levels. The main contribution of this model is that it offers insights into modelling two key aspects of human capital, health and education, and their relationship to labour supply, earnings and productivity. The model is based on Becker's (1965) household-production concept, which in turn is premised on the notion that utility is obtained not directly from market goods and services, but, rather, from final consumption goods produced from market goods and services in conjunction with one's own time. For example, leisure (a final consumption good) may be produced with the purchase of movie tickets and one's own healthy time. Some household production processes provide utility directly (e.g., the production of leisure), whereas others are inputs into other processes such as educational development or labour force participation which provide utility indirectly. A fundamental aspect of the Grossman model is that health or healthy time provides utility not only directly but also indirectly, since it is a critical input into many production processes, as described above. Accordingly, health is both a final consumption good and a capital good.

Human capital theory is premised on the notion that an increase in a person's stock of knowledge and health raises his or her productivity in both market and non-market activities. In the Grossman model, health capital differs from other forms of human capital in its effect on these activities. Health capital determines the total amount of healthy time available for them, whereas knowledge capital affects the productivity of the time spent on them. This approach suggests that health capital provides a flow of healthy time that is uniform in quality, an "all or nothing" state. An alternative formulation would be to have health capital bearing on both the quality and quantity of healthy time. Like all capital, health depreciates over time and is assumed to do so at an increasing rate with age. Consequently, investment is required to restore and/or maintain health stocks through household production activities that include inputs such as exercise, nutrition and health care. The model does not expressly include spillover effects in the production of health and education, which can be an important contributor to the efficiency of their production (e.g., the health of parents can affect child health outcomes, and some of the skills and knowledge acquired by a worker through educational pursuits can be transmitted to colleagues).

There is significant interplay between different types of human capital, specifically between education and health. In the Grossman model, higher levels of education are theorized to improve the efficiency of gross health investment. The empirical literature substantiates the existence of this relationship (Grossman and Kaestner 1997). Whether it is causal — and in which direction — is not clear, though Grossman and Kaestner conclude that the evidence suggests the existence of a pathway from education to health (i.e., individuals with higher education are better at producing health). A third variable,
time preference, may also play an intermediate role. Higher education may result in an individual placing more value on the future (i.e., it may lower an individual's time preference, which suggests that time preference is endogenous). Alternatively, a lower discount rate may encourage an individual to seek higher levels of education, which, in turn, bears on the optimal level of health capital investment.²

The Grossman model can be used to identify an individual's labour supply as a function of health. The principal implication of the model is that health is determined endogenously (i.e., by the other variables in the model rather than by exogenous/external factors). In principle, education is also determined endogenously, but since most people complete their education early in life it can be treated as an exogenous variable when modelling labour supply, whereas health capital depreciates and requires ongoing investment (Currie and Madrian 1999). As the Grossman model suggests, health is also an important aspect of human capital, and an important input into market and non-market production at the individual level.

At the aggregate level, Bloom and Canning (2000) identify four pathways by which health can affect productivity: a healthy labour force may be more productive because workers have more physical and mental energy and are absent from work less often; individuals with a longer life expectancy may choose to invest more in education and receive greater returns from their investments; with longer life expectancy, individuals may be motivated to save more for retirement, resulting in a greater accumulation of physical capital; and improvement in the survival and health of young children may provide incentives for reduced fertility and may result in an increase in labour force participation — which may, in turn, result in increased per capita income if these individuals are accommodated by the labour market.

**HISTORICAL TRENDS AND CURRENT MACROECONOMIC EVIDENCE**

The empirical literature in economic history provides substantive evidence concerning the productivity impact of increased life expectancy and reduced morbidity over the last few centuries in Europe and the United States (e.g., Costa and Steckel 1995; Fogel 1991, 1994; Steckel 2001/2002). Fogel (1991) stresses the importance of long-run dynamics and presents evidence that improvements in health which began some 300 years ago in Europe and North America have not yet fully run their course. This work suggests that an understanding of the key drivers of long-run dynamics may be of value to policy-making in developed countries even today. To this end, evidence from the historical economics literature is reviewed below, followed by empirical evidence based on data from more recent periods that makes use of the growth-accounting framework. Similar measures of health appear in both literatures.

Fogel (1991, 1994) presents historical trends in England, France and Sweden on two anthropomorphic measures associated with nutrition, namely adult height and weight (also known as body mass index). Height and weight provide different information about health. Adult height reflects the adequacy of early-childhood nutrition, whereas adult weight reflects the adequacy of adult nutrition. Using more recent evidence from Norway on the
relationship among height, weight and risk of mortality, Fogel estimates the proportion of the decline in mortality in the three countries that can be associated with changes in these anthropomorphic measures since the 18th century. Research on the relationship among height, weight and chronic disease from more recent US data provides further evidence on the detrimental health implications of below-average height and weight (relative to current North American standards). Based on this research, Fogel (1994) concludes that chronic health conditions were significantly more prevalent throughout the life cycle prior to the First World War. Consistent with Fogel’s work, Steckel (2001/2002) presents more recent evidence on anthropomorphic measures as proxies of health, supporting the notion that health can influence productivity. He found that the simple correlation between average height and log of GDP per capita ranges from 0.82 to 0.88. Furthermore, he notes that the average height of Americans is falling behind that of Northern Europeans, and that this trend may be reflective of growing income inequality in the United States.

Fogel (1994) presents evidence of the historical impact of population health on labour force productivity drawn from estimates of caloric intake in Britain and France in the 18th and early 19th centuries. He estimates that the daily caloric intake for individuals in the bottom 10 percentile of consumption in France was so low that they did not have enough energy for work, and that those in the next 10 percentile had energy for only three hours of light work (0.52 hours of heavy work). In England the situation was somewhat better. Only individuals in the bottom 3 percentile of consumption lacked enough energy for work, and those in the next 17 percentile had energy for about six hours of light work (1.09 hours of heavy work). Essentially, those in the bottom 20 percentile had such poor diets that they were excluded from the labour force. As well, many of those in the top 40 percentile were below current North American standards of average height and weight, and hence were likely subject to premature chronic conditions and mortality. Subsequent improvements in nutrition raised the energy levels of individuals in the bottom 20 percentile of consumption such that they were able to enter the labour force. These improvements in nutrition also substantially raised the capabilities of those already in the labour force. Fogel (1991, 1994) estimates that health and nutritional improvements alone can explain some 30 percent of British growth in per capita income since 1790. This value is similar to estimates of the productivity impacts of health found in cross-country studies using data from the last 50 years (World Health Organization 1999).

Recent macroeconomic research on productivity has emphasized the importance of human capital. Like physical capital, human capital in the form of education and health is durable, lasting, and subject to accumulation (Lucas 1988; Romer 1986). One approach to incorporating human capital into the macroeconomic modelling of productivity is to augment the neoclassical growth-accounting equation developed by Solow (1956). Solow’s approach to measuring multifactor productivity growth is to associate it with the residual amount of output growth not explained by growth in the key inputs of labour and physical capital. This approach is founded upon several contentious assumptions. First, it assumes that technology (which is associated with the residual) is exogenous, suggesting that labour-productivity growth rates will be the same
across economies once they reach a steady state. This runs counter to the trends of sustained differences in growth rates observed across developed countries. Another assumption of this approach is that there is perfect competition such that market prices reflect social costs — that is, there are no information asymmetries, appropriability problems, spillovers or other externalities. Intuitively, these appear to be restrictive assumptions that are likely not met in the real world, suggesting that there is a role for public policy. One can easily imagine spillover effects occurring from higher levels of human capital. They could have a substantial impact not only on an individual’s own productivity, but also on the productivity of co-workers and on society as a whole. A number of measurement issues also arise with this paradigm, the most salient of which is how to deal with technological improvements and quality changes embodied in both inputs and outputs. Quality improvements are relevant for both physical and human capital.

In response to the shortcomings of the Solow model, a new approach to growth accounting has evolved, one that attempts to model the key determinants of growth as jointly endogenous (Knowles and Owen 1997). One way to augment the Solow model is to include the accumulation of human capital as well as physical capital, while still treating technology as exogenously determined. This diminishes the importance of exogenous technological growth. The endogenous growth literature attempts to capture two aspects of the impact of health on productivity: its direct impact on the production process — for example, improvements in health can increase productivity due to reduced incapacity, disability and days off sick; and its spillover impact — for example, an improvement in the health of seniors can result in reduced personal-care time required by family caregivers who are members of the labour force.

Human capital in the form of education has received much attention in cross-country empirical growth studies, and researchers have found considerable support for its importance as a productivity driver (Harris 1999). Human capital in the form of health has received less attention, but the relatively few cross-country studies that have included some measure of health have found that it does have a significant and positive association with economic growth. Many of the empirical growth studies that include health have focused on developing countries (e.g., Bhargava et al. 2001; Hicks 1979; Wheeler 1980), though some have included a broader range of countries (Barro and Sala-i-Martin 1995; Bloom et al. 2001; Knowles and Owen 1995, 1997) and some have focused specifically on OECD countries (Knowles and Owen 1995, 1997; Rivera and Currais 1999a, 1999b).

These studies often use rather crude measures of health, likely due to the lack of data on more refined and comprehensive measures that span both a reasonable time period and a number of countries. Most studies use some measure of life expectancy or mortality (life expectancy at birth, infant mortality rates, adult survival rates), though two recent studies used per capita health-care expenditures (Rivera and Currais 1999a, 1999b). Life expectancy has increased dramatically over the post-war period in many developed countries (see Chart 1 for Canadian trends). Though mortality and life expectancy are important measures of health status, they may not capture the subtle changes in morbidity, health behaviours, health-related quality-of-life measures, or other measures of health that are particularly salient to developed countries today.
Furthermore, it is likely that the relationship between health and productivity found in these studies is driven by data from developing countries in the sample. Developed countries such as Canada and the United States are virtually indistinguishable across the health measures used (Harris 1999). Chart 2 provides evidence of this fact. As is apparent in the graph, the life-expectancy gradient is much steeper for low levels of gross national income per capita. Indeed, when Knowles and Owen (1997) estimated their specification for 22 high-income countries, they found that the health measure they used, life expectancy, was no longer significant, likely due to the lack of variability for this measure in the sub-sample. Rivera and Currais (1999a, 1999b) attempt to make a case for their use of health-care expenditures rather than life expectancy as a measure of health, but the meaningfulness of this proxy of health for developed countries is also questionable. Variations in health-care expenditures in developed countries are not highly correlated with health measures such as life expectancy and infant mortality (comparing Japan and the United States highlights this point), and it is not clear whether the marginal dollar spent on medical care reflects morbidity improvements.

A fully specified growth model should include all key inputs into the production process and all drivers of productivity, including measures of all forms of human capital; otherwise one cannot be sure whether a particular variable directly affects growth or is simply a proxy for missing factors. In particular, human capital should include a measure of education and skills, as well as health. All but one study reviewed included some measures of education and/or skills such as average years of schooling, primary/secondary/university enrolment, adult literacy or years of experience. Interestingly, the education and skills variables were not significant in most specifications in the studies, with the exception of Rivera and Currais (1999a, 1999b) and Barro and Sala-i-Martin (1995), suggesting that health may be a more important determinant of productivity, particularly for developing countries. These results might be driven by other factors, such as the nature of the proxy being used for education or measurement error in the data.

The magnitude of the coefficient for health is difficult to compare across studies, due to a number of differences in the specifications. Some studies used different measures of health, while others used comparable ones with slight variations in their specification. Some studies used the growth of labour productivity as the dependent variable, while others used the growth of total factor productivity; those that used the former generally made use of one of two denominators, an estimate of the size of the labour force or the total population. There were other specification differences as well, making direct comparison impossible.
Nonetheless, a comparison of the range of results provides a sense of the impact that health can have on productivity. To this end, Table 1 contains information on the elasticities and percentage effects of health from studies that included developed countries and from which data could be extracted. Studies by Rivera and Currais (1999a, 1999b) and Knowles and Owen (1995, 1997) suggest that between 21 and 47.5 percent of GDP growth per worker (working-age person) over the last 25 to 30 years can be explained by improvements in the health of populations (defined as health-care expenditures and life expectancy) at the country level. As noted, this range is similar to the value estimated by Fogel in his economic history work. Bloom et al. (2001) also found a significant relationship between health and GDP growth. Each extra year of life expectancy is estimated to increase a country's GDP by 4 percent.

Endogenous growth studies substantiate the importance of health for productivity growth, particularly in developing countries. Health may be equally important for growth in developed countries, but different aspects of health, such as morbidity, vitality, mental health and mental acuity, are likely more critical for these countries than increases in life expectancy. With the shift from manufacturing to services and the increasing importance of new technologies in developed countries, the human-capital needs of the labour force have changed. Intuitively, one can foresee an increasing role for mental health and acuity for knowledge workers providing high-end services. There is evidence of health improvements on this front. Recent research in the field of

![Chart 2: Gross National Income Per Capita (GNIpc) Versus Life Expectancy (LE) at Birth, 1997](chart.png)

psychology has found that populations in developed countries experienced steady gains in intellectual ability over the course of the 20th century, and some work suggests that this is attributable to improvements in health and nutrition (World Health Organization 1999). Further decreases in morbidity could also be critical for future productivity gains, given the aging of the labour force in many developed countries. Moreover, according to Mérette (2002) there is no evidence that the performance of older workers is systematically lower than that of younger workers. This may be due in part to improvements in the health of older workers and in part to a decrease in the physical demands of industries in many developed countries (as a result of both the shift towards service industries and technological advances).

### TABLE 1

<table>
<thead>
<tr>
<th>Study</th>
<th>Productivity measures</th>
<th>Health measures</th>
<th>Countries and time period</th>
<th>Elasticities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rivera and Currais (1999a)</td>
<td>log difference of GDP per worker, 1960–90</td>
<td>log percentage of GDP spent on health care</td>
<td>24 OECD countries (1960–90)</td>
<td>0.21–0.22</td>
</tr>
<tr>
<td>Rivera and Currais (1999b)</td>
<td>log difference of GDP per worker, 1960–90</td>
<td>log percentage of GDP spent on health care</td>
<td>24 OECD countries (1960–90)</td>
<td>0.28–0.33</td>
</tr>
<tr>
<td>Bhargava et al. (2001)</td>
<td>log GDP growth rate per capita</td>
<td>log of adult survival rate</td>
<td>125 countries from Pen World Tables, 107 countries from World Development Indicators (1965–90)</td>
<td>varies by GDP +ve for low-income countries -ve for high-income countries For the poorest countries, a 1% change in adult survival rate is associated with a 0.05% increase in GDP growth rate.</td>
</tr>
<tr>
<td>Knowles and Owen (1995)</td>
<td>log difference of GDP per working age person, 1960–85</td>
<td>log of (80 years less life expectancy at birth)</td>
<td>84 countries 62 developing 22 high-income (1960–85)</td>
<td>0.381 0.382 0.03</td>
</tr>
<tr>
<td>Knowles and Owen (1997)</td>
<td>log difference of GDP per working age person, 1960–85</td>
<td>log of (80 years less life expectancy at birth)</td>
<td>77 countries 55 developing (1960–85)</td>
<td>0.449 0.475</td>
</tr>
<tr>
<td>Bloom, Canning and Sevilla (2001)</td>
<td>log GDP growth rate</td>
<td>log of life expectancy at birth</td>
<td>Information not provided</td>
<td>0.04 Each extra year of life expectancy leads to an increase of 4% in GDP.</td>
</tr>
<tr>
<td>Barro and Sala-i-Martin (1995)</td>
<td>GDP growth rate per capita</td>
<td>log of life expectancy at birth</td>
<td>134 countries (developing and developed) (1965–85)</td>
<td>0.046–0.082 An estimate of .064 means that a one standard deviation increase in life expectancy (13 years) raises GDP growth rates per capita by 1.4% per year.</td>
</tr>
</tbody>
</table>
improvements in manufacturing). Nonetheless, there is a need for further research into the impact of health on productivity in developed countries using measures of health that are more salient to these countries.

**OCCUPATIONAL HEALTH AND SAFETY AND THE COST OF WORK DISABILITY**

If health capital complements firm-specific human capital in that it increases the returns to firm-specific skills and knowledge, one might expect that employers would be willing to bear the cost of investing in the health of workers in order to reap the benefits of productivity gains. But if health capital is generic rather than firm-specific, the fact that workers can take it with them from job to job suggests that firms might be unwilling to bear these costs, even if health capital increases worker productivity (Currie and Madrian 1999). In reality, health capital likely has some degree of complementarity and some generic aspects. If this is the case, firms may voluntarily invest in the health of workers but not necessarily at a socially optimal level. Consistent with this notion, developed countries have recognized the importance of labour market institutions designed to protect the health and safety of workers through financial and regulatory mechanisms. The main policy levers for providing such incentives are occupational health and safety regulation and experience-rated workers' compensation insurance. Occupational health and safety regulation covers a broad range of procedural and equipment standards and is generally enforced through a system of inspections and fines. Experience-rated workers' compensation insurance provides financial incentives for safety consciousness by varying insurance premiums at the sectoral and firm level, in an effort to tie the costs of injury and illness as closely as possible to the employers responsible for them, without unduly penalizing any one firm for costly and unpreventable accidents.

In Canada, the direct cost of work-related injuries and illnesses exceeded $5.7 billion in calendar year 2000 (Institute for Work and Health, 2002). This estimate includes indemnity payments, insurance administration expenses and medical services that are paid by employers through workers' compensation premiums (Chart 3 provides data on the growth of indemnity payments over the 1972-96 period). These direct costs substantially underestimate the true cost of productivity losses attributable to work-related injuries and illnesses. The indirect cost estimate for Canada is $12 billion. This includes costs incurred by employers to accommodate injured workers who return to work, recruitment and training costs incurred for replacing injured workers, earnings lost by workers due to injury and the lost home production of workers. Even these direct and indirect costs likely underestimate the true social cost. For instance, they do not include costs associated with pain and suffering or home care provided by family members, and the number of claims is less than the true number of work-related injuries. Clearly, the financial burden of work-related injuries and illnesses is substantial, but we do not know what proportion of this burden is preventable, the expenditures necessary to reduce the burden, or whether insurance and regulation are the most effective means of reducing the burden.

Over the past 10 to 15 years the number of work-related injury and illness claims has decreased substantially in many jurisdictions in
Canada and other developed countries (see Chart 4 for Canadian trends). In Canada, injury claims decreased by 40 percent between 1990 and 1998, despite the fact that the labour force increased by 10 percent over this period (Mustard et al. 2001). A number of explanations have been put forward for this trend. Though there is evidence to support some of these, it is not clear which factors are the predominant ones. More specifically, it is not clear what fraction of the trend is attributable to the effectiveness of insurance and regulatory mechanisms.

There is a large body of empirical literature on the effectiveness of insurance and regulatory mechanisms using econometric techniques (reviews of this evidence are provided by Curington 1988; H yatt and Thomsen 1998, Kralj 2000; and Smith 1992). Taken as a whole, the empirical evidence on the impact of regulation is mixed. The US evidence suggests that standards can reduce certain types of injuries and that a system of inspections provides, at best, modest general deterrence unless reinforced with penalties. Scholz and Gray (1997) found that regulation facilitating cooperation, such as inspections initiated by workers (regardless of penalty), can be more effective than coercive regulation such as regular inspections, unless penalties are imposed. The effectiveness of facilitative regulation is reinforced by Canadian evidence on the introduction of regulations requiring joint health and safety committees (Lewchuck et al. 1996). The evidence on workers’ compensation experience rating suggests that financial incentives can be an effective means of improving occupational health and safety. The appeal of experience rating is that it ties the cost of work-related injuries closer to the firms experiencing them, while allowing firms the flexibility to find the most efficient means of improving health and safety.

One of the challenges of regulatory and insurance mechanisms is that the nature of work and labour market experience have changed profoundly since their introduction in
Canada, yet occupational health and safety regulation and workers’ compensation continue to focus on injuries and the impact of physical and chemical exposures characteristic of manufacturing and resource-based industries. The shift away from manufacturing to services, the increasing use of technology, the introduction of new human-resource and management practices, and the growth in demand for knowledge workers have all contributed to a dramatic change in the nature of work-related injuries and illnesses (Sullivan and Frank 2000). For example, mental-stress claims more than doubled in the United States between 1980 and 1987 (Gnam 2000). Furthermore, the growth in non-standard work arrangements has not only made it difficult for workers’ compensation boards to assign firm-level responsibility for injuries and illnesses, but has also dramatically changed the nature of labour market experiences. Other factors such as income inequality, job insecurity and unemployment have also been shown to have a bearing on the health of individuals and populations (Deaton 2001; Platt et al. 1999).

HEALTH, SICKNESS ABSENCE AND FIRM-LEVEL PRACTICES

At the microeconomic level, research into the impact of health on productivity focuses on returns to employers and workers. Returns to the accumulation of health capital can be realized by employers in the form of higher profits, by workers in the form of higher wages, or by both. Pauly et al. (forthcoming) provide a theoretical framework (hereinafter the Pauly framework) to identify the principal characteristics of the production process, the market for the good or service being produced, and the labour markets that determine the size and distribution of productivity gains associated with health improvements. The framework focuses on a particular proxy for health-related productivity gains, namely decreases in sickness absence, but the premise of the model can be generalized to all health-related productivity improvements. This framework is examined below, followed by a review of the microeconomic evidence concerning the relationship between health and productivity, and employers’ efforts to capture potential health-related productivity gains through health-promotion initiatives. Consistent with this literature, particular attention is given to sickness absence as a proxy for productivity. As background for this discussion, Canadian trends in sickness absence are provided in Chart 5. Rates have been increasing for both men and women over the last few years. This may be due in part to the economic recovery during the period. Similar to work-related accident claim rates, sickness absence rates have a cyclical component (i.e., work-related accident claims tend to increase during periods of economic recovery due to factors such as the increased pace of work). Nonetheless, the rate for women has increased to levels above those of the late 1980s.

Many empirical studies on the cost of sickness absence and the cost-effectiveness of health-promotion initiatives assume that the dollar value of reducing sickness absence is simply the direct cost of wages paid to absent workers (assuming that they are paid for absences). In the Pauly framework, wages are actually the lower bound for losses. In many cases, total costs can be much higher due to indirect costs attributable to sickness absence. If the production process is team-based, or if a penalty is incurred for failure to achieve target output, then the cost of sickness absence
can be much higher if a perfect substitute is not available to replace an absent worker (i.e., if there is firm-specific knowledge capital). In the case of team production, sickness absence can also reduce the productivity of co-workers. Penalties associated with failure to achieve target output can also be significant if production is time-sensitive (e.g., perishable goods and travel/transportation services).

The benchmark case used to elaborate the framework is a single homogenous product and a simple production process in which wages reflect the incremental value of production (in which case the cost of absences is the wage rate). If health improvements are observable and transferable to a new employer (e.g., smoking cessation and weight loss), then a worker whose health improves will receive higher wage offers from competing employers. Consequently, in the long run the benefits of health improvements are fully captured by the worker in the form of higher wages.\(^\text{12}\) There is a large literature in labour economics investigating the impact of health status on wages, income and labour force participation. A thorough review of this literature is provided by Currie and Madrian (1999). In general, research supports the notion that health is associated with wages and income, though the magnitude of its impact appears to be sensitive to the measure of health used in the particular study. One of the patterns emerging from Currie and Madrian’s review is that health has a greater impact on the number of hours worked than on the wages received by workers. Even if workers do capture the benefits of health improvements in the long run, employers may still have an incentive to undertake health-promotion initiatives due to competitive pressures and in order to reap the short-run benefits, particularly if there are complementarities between health capital and firm-specific knowledge capital (Currie and Madrian 1999).

Three factors determine the degree to which the wage rate underestimates the cost of an absence: the extent to which the production process relies on team work, the size of the penalty incurred for failure to achieve target output levels, and the cost of replacing an absent worker with an equally productive one. With full employment, the wage rate is a good measure of lost output for cases in which there is a perfect substitute for a worker at the same wage rate (assuming that, in full employment, the wage rate reflects a convergence of firm, worker and societal values of labour time). The cost of absence exceeds the wage rate if the replacement worker is less productive or costs more and if the production process relies on team work or a penalty is incurred for failure to achieve target output levels.\(^\text{13}\) If there is less than full employment, the prevailing wage rate may differ from the equilibrium-wage rate, so the firm, worker
and societal values of labour time may differ as well. If this difference is small, the gains from reducing sickness absence would be similar to the full-employment case. In general, the benefits from reduced sickness absence are greater than the wage rate. It should be noted that the impact of sickness absence on productivity is greater when measured by output per worker than output per hour, since some of the output not produced due to sickness absence is offset by the reduced number of hours worked.

The Pauly framework underscores the potential productivity gains of reducing sickness absence through investments in health capital. Implicit in this line of thinking is the assumption that poor health is the principal reason for sickness absence. But the etiology of sickness absence is quite complex; poor health is but one of many factors that have a bearing on sickness absence. Examples of other factors include personality, job-related attitudes and social context. Two recent literature reviews (Alexanderson 1998; Harrison and Martocchio 1998) identify a broad range of causal factors investigated in the empirical literature on sickness absence and find that few studies have investigated the impact of health on sickness absence. In particular, short-term health conditions are rarely studied. Nicholson and Martocchio (1995) describe this gap in the literature as a "black hole." Alexanderson notes that even though most studies in her review are from the fields of epidemiology and medicine, most do not use a medical model or even consider health status as an explanatory variable. This lack of attention to health is likely due to a focus on sources of variance that are perceived to be avoidable (and thus amenable to change), though the distinction between avoidable and unavoidable is blurry at best. For example, some people may believe that health status is predetermined and that changes in health status are unavoidable, but an intervention such as an influenza shot can significantly reduce the incidence of colds and flu, and thus may reduce sickness absence more readily than an intervention designed to increase workers' engagement in their work.

Nonetheless, there is evidence to support the notion that health status has a strong influence on sickness absence. Most empirical studies of sickness absence that have included measures of long-term health, chronic conditions or health behaviours have focused on self-reported health status, smoking, illicit-substance use and alcohol consumption. The Whitehall II study provides some of the most compelling evidence for the impact of these factors on sickness absence (Marmot et al. 1995; Marmot et al. 1993; North et al. 1993). Reported "average" or "worse" health over the 12 months preceding the survey was associated with significantly higher levels of sickness absence compared to reported "good" health — a 60 percent increase in short spells (seven days or less) and a twofold increase in long spells (more than seven days). Significantly higher levels of sickness absence were also observed for individuals reporting recurrent health problems, long-standing illnesses or psychiatric symptoms. Mental health factors such as depression, anxiety and emotional stress are a frequently reported cause of sickness absence, particularly among women (Stansfeld et al. 1995), and have been found to be significant predictors of absence and disability (Garrison and Eaton 1992; Kessler et al. 1999; Kouzis and Eaton 1994; Simon et al. 2001; Skodol et al. 1994). Studies of health-related behaviours have found that smokers have higher rates of absence than non-smokers (Bush and Wooden 1995; Leigh
1995; North et al. 1993; Parkes 1987) and that illicit-substance users have high rates of absence (Bass et al. 1996; Normand et al. 1990). The impact of alcohol consumption is more complex. Problem drinking (a level of alcohol consumption that results in social dysfunction) appears to be the point at which sickness absence is affected (Beaumont and Hyman 1987; Casswell et al. 1988), causing long-term absences in particular (Marmot et al. 1993).

Acute conditions related to respiratory and gastrointestinal conditions are the primary reasons for short-term absences (Stansfeld et al. 1995), yet, as noted, few studies have investigated the nature and impact of short-term health conditions on sickness absence. This is surprising given that more immediate inroads into reducing sickness absence might be made by addressing the factors that cause acute conditions. For example, Nichol et al. (1995) found a 43 percent drop in the rate of cold- and flu-related sickness absences among adults receiving an influenza vaccine instead of a placebo.

The evidence regarding the impact of health status on sickness absence, though preliminary, suggests that firm-level initiatives such as health promotion are one means by which employers can reduce absence and increase productivity. Employers have undertaken a variety of initiatives, ranging from targeted to multi-component programs (fitness/exercise programs are the most widespread). A large empirical literature on the impact of these initiatives on health and productivity has accumulated over the past 30 years. Most of the empirical studies have considered one or more of five outcomes: risk reduction or behavioural change, health/medical-care costs, sickness absence, turnover and other proxies of productivity. One of the motivations for these initiatives in the United States was the growing cost of employee health care in the 1970s and 1980s, which was increasing at a rate of up to 30 percent per year (Conrad 1988). The focus on health-care costs in many US employer initiatives is understandable given that health insurance is generally provided by employers and can make up a substantial component of the benefits provided to workers. In Canada, it is a less salient cost for employers, since health insurance is primarily funded by the public sector, but is nonetheless a factor to consider at the macro level. Canada has also experienced substantial increases in health-care costs since the 1970s, and there is evidence that a growing proportion is paid by the private sector (Polanyi et al. 2000).

Studies of the impact of employer-sponsored health-promotion initiatives generally use quasi-experimental methods (participants self-select into the program) or non-experimental methods (no control group — e.g., before/after comparison) and cost-benefit or cost-effectiveness analysis. The dearth of rigorous research, particularly in the earlier literature in this field, has been commented on in most reviews (Aldana 1998; Baun 1995; Fielding 1990; Heaney and Goetzel 1997; Messer and Stone 1995; O'Donnell 1997; Pelletier 1991, 1993, 1996, 1999, 2001; Saha et al. 1994; Shepherd 1992; Warner 1992; Warner et al. 1988). As a whole, these studies show mixed evidence, with non-experimental studies generally demonstrating positive results and more rigorous studies demonstrating less positive results. On average, studies using experimental designs had positive results approximately 25 percent of the time, quasi-experimental designs 50 percent of the time and non-experimental designs 100 percent of the time (Heaney and Goetzel 1997; O'Donnell 1997). In terms of sickness absence, the evidence on the impact of health promotion is mixed (Baun 1995).
range of studies reviewed by Shephard (1992) can be accepted as the true program effects, then the impact on sickness absence is positive but modest; most of the studies found effect sizes in the range of 0.5 to two days per year of improved attendance. In Messer and Stone's (1995) review of cost-benefit studies, those that included the cost of reduced sickness absence as a benefit found that benefits exceeded costs in all cases (benefit-cost ratios ranged from 1.07 to 3.90); for some of the studies, the positive effects may also have been driven by reduced health-care costs.

The quality of research has improved substantially in the most recent generation of studies, providing a rationale for employers to consider such programs as a means of reducing the health and economic costs of illness. The findings provide some support for the hypothesis that health and fitness can have an impact on sickness absence and productivity, though the durability of these effects has not been established. A realistic assumption is that a program will need to engage workers on an ongoing basis if the changes are to be sustained. Though most studies have found only modest reductions in sickness absences as a result of health-promotion programs, the direct costs of these absences represent only a lower bound for the costs attributable to sickness absence, suggesting that productivity gains may be higher.

Polanyi et al. (2000) are less optimistic about the potential for firm-level health-promotion initiatives. They list five reasons why such initiatives may be limited in their ability to improve worker health: (1) this approach does not address the sources of human motivation and behaviour that bear on health, namely the social and economic determinants of health; (2) lifestyle is a less important factor bearing on health than socioeconomic status (Marmot et al. 1991, 1993), which has health effects that can endure even after retirement (Wilfson et al. 1990); (3) if behavioural and lifestyle changes are to be enduring, the social and cultural context that engendered them must be modified as well; (4) health-promotion initiatives reach only a limited number of workers, since they tend to be implemented in white-collar settings; and (5) unless other job-related and organizational factors are also addressed, workers may perceive such initiatives as a self-serving effort by firms to reduce their health-care and absenteeism costs.

Indeed, it is becoming increasingly evident that general health and functioning are very much affected by work experiences. There is a growing body of evidence showing that psychosocial workplace factors such as job control, psychological demands and social support have an important bearing on workers' health (Shannon et al. 2001). This suggests that broadly based organizational initiatives may be more successful in improving workers' mental and physical health and productivity than focused health-promotion programs. Such broad-based initiatives can include a range of elements such as: redesigning worksites with ergonomic principles in mind; redesigning work flows and communications channels to enhance communication and social support; and providing flexible work hours, leave programs and daycare facilities. Unfortunately, there are few experimental studies in the literature investigating the productivity impact of such initiatives.

**SUMMARY AND CONCLUSIONS**

Economic well-being in developed countries is growing increasingly dependent on international markets and integrated trade. To
maintain a high standard of living in the global marketplace, countries must remain competitive, which in turn depends on the fostering of an innovative, productive labour force. Researchers and policy-makers are becoming increasingly aware that a country’s ability to innovate and remain productive depends on the characteristics and quality of its human capital, key elements of which are education, skills and health. Most recently, the links between population health and economic productivity have become a significant policy concern. With the aging of the labour force in Canada and many other developed countries, labour force health will become an even more important issue in the near future. On the positive side, there is evidence that older workers in developed countries are no less productive than their younger counterparts.

Fogel’s research in economic history highlights the significance of population health for productivity growth. He provides compelling evidence for the important role of nutrition, particularly in early childhood, on health and functioning throughout the life cycle. Fogel points out that these historic trends still have a bearing on the health of populations in developed countries today. This research provides valuable insights into the critical role that policy can play in supporting population health and ultimately productivity. For example, the enduring effect of childhood experiences is one of the themes in this work, suggesting that financial support for low-income families, parental leave policies and child-care policies can help to ensure healthy child and adult outcomes. The evidence from macroeconomic studies on health and productivity corroborates the evidence from economic history. Health is indeed an important driver of productivity even today. This area of research focuses on life expectancy as a measure of health, which can be an important productivity driver for developing countries but is less salient for developed countries. In order to better serve the information needs of policy arenas in developed countries, future research in this area should identify more refined measures of health to incorporate into their empirical analyses (e.g., measures of functional status and mental health).

Work-related injuries and illnesses are a major source of productivity losses for society as a whole, suggesting that public policy can play a crucial role in providing incentives for employers to dedicate more resources to occupational health and safety. Most developed countries recognize the importance of labour market institutions designed to protect the health and safety of workers, and have established financial and regulatory mechanisms for this purpose. Different policy levers are available under the umbrella of regulation and insurance, some of which are more effective than others. The evidence suggests that occupational health and safety regulation has had only a modest impact, though preliminary evidence shows that regulation focusing on facilitation rather than coercion may be more promising. The evidence for experience-rating suggests that it is effective, partly because financial incentives allow employers the flexibility to identify the most efficient means by which to improve workplace health and safety. Both insurance and regulation tend to create incentives that focus on acute injuries and the health impacts of physical and chemical exposures characteristic of the industrial sector. Consequently, many of the important factors affecting labour force and population health in developed countries remain outside the purview of traditional...
insurance and regulatory domains, due to the changing nature of work and labour market experiences in these countries.

Sickness absence is an easily measured proxy for productivity that is often investigated in the organizational practices literature. Yet, surprisingly, the empirical literature on the causes of sickness absence has given little attention to the role of health status as an explanatory variable. The work that has been undertaken in this area suggests that chronic and acute physical and mental conditions, as well as health-related behaviours, explain a significant portion of sickness absence. This evidence provides support for health-promotion initiatives as a means of reducing sickness absence and increasing productivity. Empirical studies evaluating workplace health-promotion programs find that they are effective in reducing absence and health-care costs, though these reductions are modest and the durability of program effects is not known. Nevertheless, the direct costs of absences are only a lower bound for productivity losses attributable to them; the productivity gains to be realized from decreasing sickness absences may be substantially higher. Future research in this area should focus on developing the tools to quantify the various indirect costs associated with sickness absence, as well as improving the methodological rigour of intervention studies. Furthermore, there is evidence to suggest that such initiatives should consider a broader set of organizational factors related to work and the workplace that bear on the health and productivity of workers instead of narrowly focusing on behaviour and lifestyle.

The fact that firm-level initiatives appear to have had a limited impact on productivity suggests that the public sector has a role to play in improving the health of the labour force and population as a whole and, in turn, overall productivity. Identifying priorities in order to achieve the greatest gains from the resources invested is a difficult task at best. Certainly, increasing health-care spending will not necessarily result in higher levels of population health, as evidenced by the differences in per capita health-care expenditures and health profiles across OECD countries. The multifaceted nature of the factors that influence health suggests that policies in a number of areas traditionally considered outside the purview of health policy may be important avenues by which the public sector can have an impact on population health. Key areas are labour market policy, education policy, and child-care and parental leave policy. Furthermore, though improving population health is an important societal objective, there are many other objectives competing for scarce public resources. Achieving an optimal balance when addressing societal objectives requires a sound understanding of the policy options, their impact on the various objectives, and the costs associated with each.

NOTES

1. Muurinen (1982) provides an alternative formulation in which education is theorized to lower the rate of depreciation of health stock.

2. A lower discount rate may also encourage investment in health directly.

3. There are other measurement issues with regard to capturing the full impact of health on welfare. In particular, utility is derived directly from health, as well as indirectly through its role in market and non-market production activities, whereas only market production activities are captured in standard output measures. The direct value of health may be captured to some degree through market goods and services purchased to improve health. The indirect value of health from non-market activities may also be captured to some degree in this way. At the core of this measurement issue is the fact that standard measures of output do not capture non-market resources and activities. Costa and Steckel (1995)
The Impact of Health on Productivity: Empirical Evidence and Policy Implications

compare alternative methodologies of measuring welfare changes arising from changes in the health of populations as an illustrative exercise to assess the consistency of rankings.

4 Significance does not necessarily imply causality. In fact, the direction of effects might run from productivity growth to health, rather than the reverse. All of the studies address the reverse causality issue by testing for potential endogeneity.

5 Currie and Madrian (1999) classify health measures typically available in the data sets of developed countries into eight categories: self-reported health status, presence of functional limitations on the ability to work, presence of functional limitations on other activities, presence of chronic/acute health conditions, health-care utilization, clinical assessments of health, nutritional status and mortality.

6 GDP per capita growth is highly correlated with GDP per worker growth, with the latter driving the former. None of the studies corrected for the fact that average hours of work can vary substantially from country to country.

7 Leigh et al. (2001) estimate that the total cost of health care and lost productivity due to occupational injuries and illnesses in California is on par with the costs of all cancers combined, and comparable to the costs of heart disease and stroke.

8 These explanations support one of two propositions: that the trend reflects real decreases in work disability, or that the trend is a reporting phenomenon. Following are some of the explanations offered: de-industrialization has resulted in a greater proportion of the labour force being employed in the service sector, which is inherently safer; older capital is being replaced with new capital, which embodies ergonomic and technological improvements that are inherently safer; employers are doing a better job of instituting safety measures in response to insurance and regulatory incentives or because of an increased awareness of the value of health capital; employers are adopting more aggressive claims-management practices; and workers’ compensation boards have tightened their eligibility requirements.

9 This study also found that firms that are reluctant to form joint health and safety committees show poorer health and safety performance, suggesting that the climate of the internal-responsibility system is an important element in its success. The authors suggest that, if the internal-responsibility system is to be effective, special measures may be required to educate employers and workers in such cases.

10 Mental-stress claims still represent only a small fraction of workers’ compensation claims in North America, though this may be due in part to legislation passed in the early 1990s in many jurisdictions limiting compensability for such claims (Gnam 2000).

11 Decreases in sickness absence are frequently used as a proxy for health-related productivity gains attributable to employer-sponsored initiatives, primarily because sickness absence is readily observable and measurable, and data regularly collected by human-resource departments. Two other proxies considered in the literature are disability and turnover. It should be noted that health improvements can increase not only the amount of time available for work, but also the quality of time spent at work. More comprehensive empirical analyses attempt to assess the entire range of direct and indirect benefits and costs associated with a health improvement.

12 In the case of a salaried worker who is required to make up lost production time, the benefits are in the form of more convenient hours. Note that the value of lost leisure time is not captured by traditional output measures but is nonetheless an important aspect of social welfare.

13 Firms may attempt to insure against losses by hiring extra workers as backups to cover for absences. This imperfect remedy will be relatively less costly for larger firms than for smaller ones.

14 One of the shortcomings of the model is its narrow focus on sickness absence, rather than the impact of worker health on productivity in general. However, the basic tenets of the model can be generalized to this broader perspective.

15 Harrison and Martocchio (1998) review over 500 empirical studies on sickness absence culled from a variety of disciplines conducted from 1977 to 1996.

16 Frequently cited methodological shortcomings are failure to use a control group or randomization of individuals between program and control groups, failure to adjust for confounders, lack of standardized measures used for exposure and outcome, short time period of studies, small sample sizes, use of self-reported measures, failure to consider all indirect costs and benefits associated with an initiative, bias introduced by evaluations being performed by program advocates, and failure to consider the worksite as the unit of analysis when initiatives are implemented at a selection of worksites.

17 Baun (1995) reviews studies that assess a range of health-promotion initiatives for their impact in terms of reducing sickness absence. He concludes that the evidence for the effectiveness of smoking-cessation programs, health-risk assessment programs, and exercise programs is mixed. Only stress-reduction programs appear to have a powerful effect on sickness absence.
REFERENCES


The Impact of Health on Productivity: Empirical Evidence and Policy Implications


