

NBER WORKING PAPER SERIES

WORK AND THE DISABILITY TRANSITION
IN 20TH CENTURY AMERICA

Sven Wilson
Joseph Burton
Benjamin Howell

Working Paper 11036
<http://www.nber.org/papers/w11036>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2005

We thank the National Institute on Aging (AG10120) for financial assistance. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

© 2005 by Sven Wilson, Joseph Burton, and Benjamin Howell. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Work and the Disability Transition in 20th Century America
Sven Wilson, Joseph Burton, and Benjamin Howell
NBER Working Paper No. 11036
January 2005
JEL No. I12

ABSTRACT

Using data from Union Army pensioners and from the National Health Interview Surveys, we estimate that work-disability among white males aged 45-64 was 3.5 times as high in the late 19th century than at the end of the 20th century, including a decline and flattening of the age-profile since 1970. We present a descriptive model of disability that can account for a) the secular decline in prevalence; b) changes in slope of the age-profile; and c) periods of increasing prevalence. The high level and relatively flat slope of the historical disability age-profile is consistent with the early onset of chronic conditions and with high mortality associated with a subset of those conditions. We show that many common conditions in the 19th century have been either eliminated, delayed to later ages, or rendered less disabling by treatment innovations and the transformation of the workplace. These improvements have swamped the effect of declining mortality, which put upward pressure on disability prevalence. Given the low rate of mortality prior to age 65, technological changes will likely induce further reductions in work-disability, though recent increases in the prevalence of asthma and obesity may eventually work against this trend.

Sven Wilson
Department of Political Science and Economics
Brigham Young University
Provo, UT 84602
sven_wilson@byu.edu

Joseph Burton
University of Chicago
jburton@cpe.uchicago.edu

Benjamin Howell
Brigham Young University
b_howell@byu.edu

1. Introduction

Understanding the social consequences of rapid population aging is one of the central questions of our day. Over the past century mortality has fallen at all ages, leaving us with the question of what the coming era holds in terms of quality of life for an increasing number of people. Addressing this challenging question is complicated by considerable uncertainty and controversy associated with current trends in a variety of important health measures.² Perhaps the central question guiding the debate is whether increases in life expectancy will result in a population that lives longer but with a higher level of disease and disability. Currently, the baby boom generation is aged 40-60 and the ability of our private and public resources to accommodate the graying of the baby boomers depends critically on the ability and propensity of this generation to work and upon the health care resources they will demand in the next two decades.

The recent literature on disability is limited in two important ways that we want to address in this analysis. First, almost all the attention has been on disability among the elderly (age 65+). Second, most of the literature concentrates on disabilities that affect “activities of daily living” or ADLs. These include difficulties in activities such as bathing, dressing or toileting. In most cases, people with ADL disabilities are quite severely disabled. But many people (particularly the non-elderly) have disabilities that affect their lives in socially important ways—such as the types of jobs they are able to have or the timing of retirement from the labor force—even though they do not have ADL disability. In sum, the recent literature informs

² A lengthy literature exists on this topic. Important studies in the last decade include Crimmins (1996); Crimmins & Saito (2000); Crimmins, Saito & Reynolds (1997); Cutler (2001); Freedman & Martin (1998); Freedman, Martin and Schoeni (2002); Manton & Gu (2001); Manton, Corder & Stallard (1997); Ostechegea, et al. (2000); and Waidman Bound & Schoenbaum (1995).

questions of the need for long-term care of the elderly, but has relatively little to contribute to questions regarding labor market activity.

Researchers' preoccupation with older-age ADL-disability reflects not only the policy importance of the health of the elderly, but it reflects a profound shift that has occurred over the past century concerning the extent and nature of disability—what we refer to here as the disability transition. A century ago disability was *not* a phenomenon predominantly of the elderly. Debilitating conditions struck early and were often of lengthy duration. Furthermore, physical demands of work and other life activities were very high. These factors resulted, we will show, in an age-profile of disability among younger adults that was strikingly high and relatively flat. Today, even though many more people are living to advanced ages, we have pushed back both the onset and the debilitating effects of chronic illnesses.

Over twenty years ago Fries (1980) predicted that morbidity would be compressed increasingly into the later years of life. The “compression of morbidity” hypotheses has turned out to be perhaps only half right. The onset of morbidity and disability has definitely been extended, but Fries assumed an upper bound to life that is still elusive, as disability among the elderly seems to be declining across age groups even while life expectancy continues to grow. Furthermore, Fries' hypothesis cannot speak to the possibility of increasing disability during certain transitional periods. The simple model we present later not only captures the predictions of Fries but can also account for short-run periods of increasing disability even as longer-term trends imply disability decline.

In this paper we outline the key ideas of the disability transition and summarize evidence supporting it, including new evidence we present on the importance of work disability. Our empirical focus is on two key elements of the transition: 1) the substantial decline in work

disability among adults aged 45-64 and 2) a significant shift in the types of chronic illnesses associated with work disability. It is our conjecture that a historical focus on early disability may provide valuable information in the understanding of disability over the life course, since several hypotheses are possible concerning how the age profile of disability should have evolved over the past century, all of which are consistent with declining disability past age 65, but which differ markedly in terms of disability patterns at younger ages.

2. The Disability Transition

2.1. The Disablement Process

Health is a topic studied across a multitude of both biological and social sciences. As a result, researchers often disagree over how to define disability and how to distinguish it from other health variables. Disability, as defined by the World Health Organization (WHO), is “the inability to perform socially prescribed roles due to medically defined conditions, impairments, and functional limitations (1980).” Disability differs from other health measures in that it depends fundamentally upon the individual’s environmental and social context, including socioeconomic status, employment, human capital and family and social networks. We can never identify or understand trends in disability by divorcing environmental and social processes from biological and medical processes. The concept of disability only makes sense when it is embedded in a social and environmental context.

Scholars in recent years have become more careful and precise about distinguishing between different health concepts, including disability. Building on the work of Nagi (1965, 1969, 1991), and the National Institute of Medicine (Pope & Tarlov, 1991), Verbrugge and Jette (1994) articulated a highly influential, sociomedical model of the disablement process that is useful in epidemiological, demographic and clinical research and that has become the standard in

the field. The disablement process is highlighted in Figure 1 below. There are four main conceptual variables. The first is disease (either chronic or acute) and injuries. Disease leads to “impairment,” which is actual damage to body tissues, such as when arthritis causes tissues in the joints to become inflamed or when emphysema kills tissue in the lungs. Impairments, then, lead to functional limitations, which constitute the body’s reduced ability to perform particular physical or cognitive tasks, such as climbing stairs, bending, sitting for long periods, or picking up coins off a table surface. Reduced lung capacity, for example, may cause a difficulty in walking short distances, and swelling in the joints may affect the ability to pick up small objects with one’s fingers.

The final stage of the model is disability, which is the inability to perform some socially prescribed role, such as work, taking care of the house, or conducting personal hygiene. Functional limitations are frequently confused with the disability, but they are not the same thing. Whether or not fine motor dexterity is disabling, for instance, depends on the individual’s perceptions of her limitations and her social context. An artist or a surgeon may find a lack of fine motor skills as highly disabling, while a salesperson may face no disability. From the perspective of economic theory, even if we assume that the biomedical and social environments are exogenous, disability will remain fundamentally an endogenous, equilibrium concept, since individuals have some ability to choose and shape their environments.

Our focus here is on upon the role of variables that are included under the broad categories of “medical technology” and “environmental/social factors.” Figure 1 illustrates that each linkage in the disability pathway is a function (F_1, F_2, F_3, F_4) of these variables and the variables that lie on the health pathway. The initial determination of disease, F_1 , is also a function of genetics, since most pathologies result from a combination of genetic and non-

genetic risk factors. Medical technology, which would include immunizations, can also work to prevent the development of disease.

Progression along the disablement pathway can be influenced at every point by medical and environmental/social factors. Take, for example, the case of asthma, a disease whose rising prevalence has caused renewed public health concern over the past decade. Whether or not an asthmatic person becomes disabled depends not only upon the medical treatment he/she receives, but upon factors such as the local climate and the level of physical activity. Similarly, whether

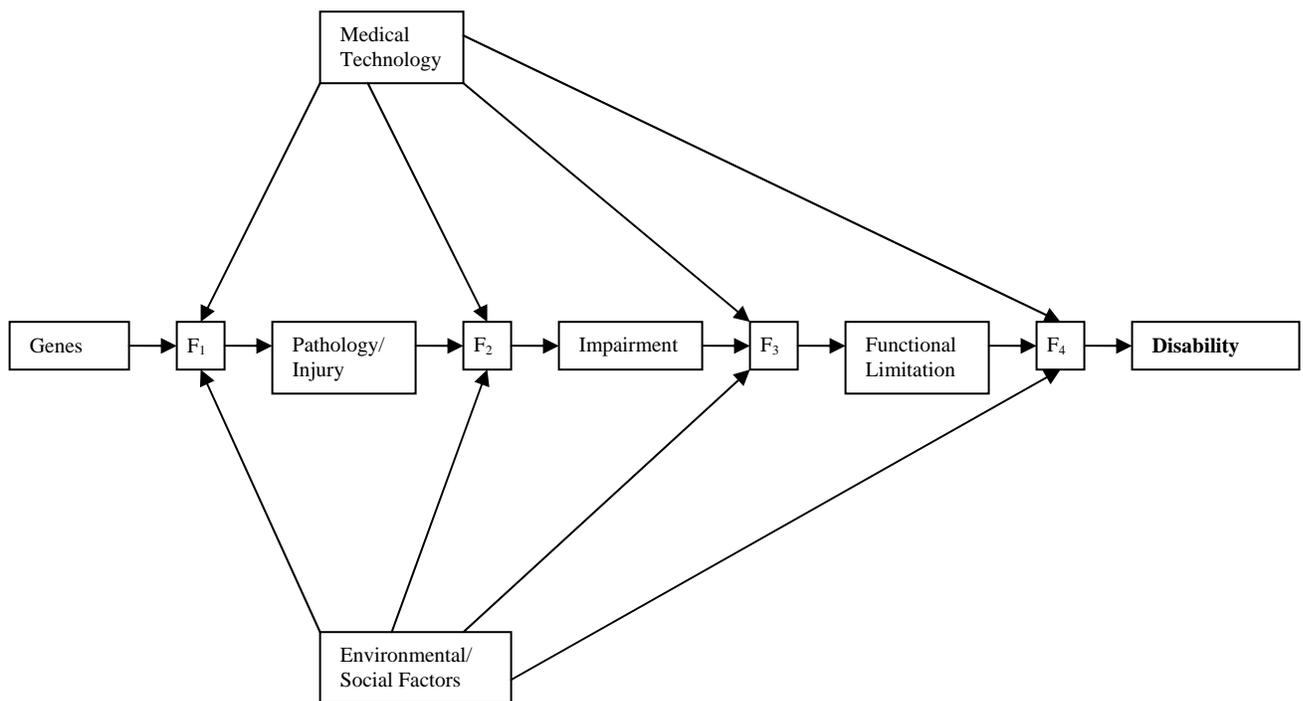


Figure 1: The Disablement Process

or not arthritis becomes disabling depends on medications and other treatments and on the physical demands of life that the arthritic individual may face.

An important aspect of disability not captured in the scheme of Figure 1 is the severity of disability. Arthritis, for example, can become completely debilitating or it may result in nothing more than a minor inconvenience. Thinking carefully about severity is particularly important from empirical and policy perspectives. The researcher's view of severity will, in part, determine the types of disability and other health measures employed in a given analysis. Recently, the primary focus in the disability literature has been the activities of daily living (ADLs), such as being able to bathe, use the toilet, or get in and out of bed, and instrumental activities of daily living (IADLs), which are cognitive tasks such as being able to balance a checkbook or managing medications.

Unfortunately, a portion of the disability literature implicitly assumes that ADLs and IADLs capture the entire set of potential disabilities.³ ADL-disability is the most severe type of disability present in the population, and identifying trends in ADL-disability is an important and policy-relevant task. But individuals can be disabled, as discussed above, without having these severe disability outcomes. We argue that a significant portion of work disability is serious enough to affect important health-related decisions (such as whether to work, where to live, and how often to seek medical care), even though severe, ADL-disability may or may not exist.

2.2. Disability and the historical record

Given vast improvements in medical technology, public health reforms, superior nutrition, and in reduced physical demands of work, it seems probable that population health—whether in terms of disease, disability or mortality—should have seen vast improvements over the past century. Even though it is safe to say that populations in the developed world are much healthier than they were a century ago, this is not to say that all “improvements” in

environmental and social variables or in medical technologies have had positive effects on disability trends as a whole or, in particular, on the individual stages of the disablement process discussed above.

Historical change in a variety of domains has wrought changes in health-related outcomes over the past century. Improvements in medical technology, particularly the development of antibiotics and immunizations against serious infectious diseases, have affected the first stage, F_1 , which is the development of pathology and injury. Furthermore a vast array of treatment options from cardiac surgery to arthritis pain medication, has mitigated the development of impairments and functional limitations. Rehabilitative and adaptive devices such as advanced wheelchairs and work-place technologies for the disabled also would play a role in weakening the link between functional limitation and disability. Extensive public health programs that have been instituted over the past century—clean water, improved sanitation, pollution regulation, and health education programs, to name a few—also played important roles in stopping the spread of disease. However, some public health advancements may have served to accelerate progress along the disablement pathway for some injuries and pathologies. Side effects from treatment may have cured or accommodated one disease but resulted in new pathologies, injuries, or impairments.

Another feature of medical advancement is that various aspects of the human condition that were previously thought to be of a moral nature—particularly psychological disorders and substance abuse—are now commonly seen as diseases. Instead of viewing alcoholism, for instance, as a moral failing, much of society (with the assistance of the medical community) is likely to now see it as a disease that can have debilitating consequences. Similarly, common

³ For instance, the most thorough review of the recent trends in disability to date (Freedman, Martin & Schoeni, 2002) classify “any disability” as either ADL or IADL disability.

psychological ailments, such as depression or anxiety, are now seen as health problems. These conditions are often treated medically, and doctors, patients and the community at large is comfortable attributing the consequences of these diseases, such as the inability to maintain employment, as legitimate disabilities.

The industrial revolution and the transformation of the American economy likely played a very important role in the disability transition, though we do not accumulate direct labor market evidence in this paper. The conditions of the workplace play a fundamental role in the final stage of the disablement process, F_4 . Costa (2002) argued that the structural shift towards accounted for only about 10% of the decline in the functional limitations in the twentieth century, but this refers only to changes across jobs, not the changing of work *within* jobs. The industrial revolution of the late nineteenth century resulted in the mechanization of manufacturing, agriculture, and even the trades and professions, and it is safe to say that almost all jobs experienced technological innovations that made them physically less demanding and often safer. Furthermore, average hours worked declined significantly over the century, decreasing the physical demands of labor.

But not all aspects of increasing industrialization were positive. Industrialization resulted in high rates of pollution and toxic waste, both in the workplace and the larger community, which persisted well into the twentieth century. The health effects of industrialization are poorly understood (mostly because of a lack of micro data on individual-level health), but some evidence points to the importance of pollution. The industrial revolution also brought rapid urbanization and immigration. These forces, coupled with a high rate of internal migration, brought people into closer association and with higher frequency than ever before, which facilitated the spread of infectious disease. It is possible that population health was deteriorating

significantly in the late 1890s and early 20th century, implying that disability measured at a slightly later point than 1893 would yield even greater rates of disability. For instance, Wilson (2003) showed that the prevalence of respiratory disease was rising between 1895 and 1910, which is consistent with rising levels of pollution and population density.

The twentieth century also saw vast changes in consumption patterns of ordinary Americans. Although we would expect individuals to purchase better health as their income rises, we have seen many negative patterns as well. Smoking, as evidenced by aggregate cigarette production, increased rapidly through the first half of the twentieth century (Wilson 2003). In recent decades smoking rates have fallen, but obesity has skyrocketed. Reductions in the physical demands of labor may have reduced some types of disabilities, but the associated sedentary lifestyle (along with sharp declines in the real cost of food) may have had even more serious long run consequences in terms of health and disability.

Finally, disability is fundamentally a psychological concept that depends not only on biomedical determinants but also on how an individual interacts with and interprets his or her social context. As incomes have risen dramatically in developed countries in the last century, aspirations have risen as well. Just as “relative income” may explain consumption behavior, “relative health” may keep disability rates (especially general measures such as work disability) higher than we would expect given a decline in functional limitations. To the extent that disability depends on one’s changing level of aspirations (driven by observations of the well-being of one’s peers), disability rates may be strongly persistent or even increasing. Changes in society’s aspirations regarding disability may have contributed to both large reductions in the labor force participation rate of men under 65 in recent decades and in increases in enrollment in disability compensation programs such as SSDI.

In summary, four broad lessons can be drawn from this brief and informal analysis of the historical record. First many important social and economic trends have produced both positive and negative impacts on the progression from disease to disability, the net effect of which remains uncertain in some cases. Second, because of the presence of both positive and negative consequences, the periods used to make historical comparisons may prove to be critical, since levels of disability are likely to have fallen and risen at different points in time (and for different ages). Third, social trends which may weaken one link in the disablement process may strengthen one of the other links. Fourth, the complex relationship between disease and disability is likely to change significantly over time, and that change may be reflected at different stages of the disablement process. For instance, sometimes medical advancements may result in a suppression of the risk of acquiring a given disease, while others may reduce the incidence of functional limitations associated with the impairments caused by the disease. Consequently, it is important to gather and analyze information on all the stages of the disablement process—disease, impairment, functional limitation and disability—to more completely understand the dynamics of population health. Because disability is a contextually-dependent variable, declines in diseases, impairments or functional limitations are not guaranteed to result in declines in disability.

2.3. Modeling the age structure of disability

Declines in disability at older ages are consistent with a variety of plausible shifts in age-disability profiles across time. To guide our thinking on this issue and capture the salient aspects of the historical record just discussed, consider the following descriptive model of disability prevalence that simplifies the model outlined in Figure 1 and that corresponds with the historical data we are using in our empirical analysis. We note at the outset that this is not a behavioral

model; it describes the age structure of disability prevalence, not the important choices individuals make in regards to their health, labor supply or their environments.

First, take a birth cohort of size N_a and assume, initially, that there is no mortality. We take as a starting point the existence of a potentially debilitating chronic condition represented by a dichotomous variable: $C=1$ if any condition exists, and $C=0$ if not.⁴ The probability of a member of the birth cohort having a condition at age a is $p_a = \text{prob}(C=1|a, \theta)$. As shown, this probability depends on age and a variety of social and environmental factors, which we refer to as θ , that vary across individuals and across time. The probability function is also dependent on the Functions F_1, F_2, F_3 described earlier, but these are not explicitly modeled here. Similarly the presence of a disability is defined as $D=1$ if disabled and $D=0$ if not, and the probability of the *potentially* debilitating condition *actually* resulting in a disability is $r_a = \text{prob}(D=1 | C=1, a, \theta)$ (naturally, if $C=0$, then $D=0$ as well).⁵

The next step is to account for mortality. An individual's status at age a is either alive ($A=1$) or dead ($A=0$). The probability of being alive at age a , given a condition, is $s_{ac} = \text{prob}(A=1 | C=1; a, \theta)$, and the probability of being alive in the absence of a condition is $s_{ah} = \text{prob}(A=1 | C=0; a, \theta)$. A further refinement would be to allow disability to affect mortality above and beyond the effect induced by the debilitating condition, but for the sake of parsimony, we forego consideration of that issue.

Finally, the discussion above makes no explicit consideration for acute conditions and accidents, some of which may be potentially disabling, and some which may be fatal. We divide

⁴ Obviously there are a variety of conditions that can be debilitating, and the likelihood of becoming disabled depends on the combinations of conditions an individual has and how that set of conditions interacts with the social context of the individual.

⁵ As discussed earlier, a variety of biomedical, social, and environmental factors will affect the probability of a disability occurring. For the sake of analytical convenience, we do not model transitions out of disability into non-disability, though we recognize the importance of such transitions in modern times.

acute episodes into two types. The first are those which may or may not be fatal but which do not have debilitating effects in the long-run. Blood loss from an accident would be an example. People who have had this type of acute event have the same survival curve, s_{ah} , as people who have no condition. The other type of event is that which either raises the probability of acquiring an acute condition or lowers the survival curve for the person below s_{ah} . For instance, rheumatic fever is an acute condition that often causes valvular heart disease. In this case, we assign a value of $C=1$, meaning we treat the individual as if he had had a chronic condition, and we assume he is subject to the s_{ac} mortality schedule.

Given this framework, we write the number of disabled persons as

$$D_a = p_a r_a s_{ac} N_a,$$

while the number of health people, H_a , and the number of people at risk for a disability because of a condition, R_a , are:

$$H_a = (1 - p_a) s_{ah} N_a$$

$$R_a = p_a (1 - r_a) s_{ac} N_a.$$

Therefore, the age-specific disability prevalence rate, ρ_{Da} , after simplifying terms, is:

$$\rho_{Da} = \frac{D_a}{D_a + R_a + H_a} = \frac{p_a r_a s_{ac}}{p_a s_{ac} + (1 - p_a) s_{ah}}$$

The marginal impact of each variable on prevalence is:

$$\frac{\partial \rho_{Da}}{\partial p_a} = \frac{r_a s_{ac} s_{ah}}{(p_a s_{ac} + (1 - p_a) s_{ah})^2} > 0,$$

$$\frac{\partial \rho_{Da}}{\partial r_a} = \frac{s_{ac} p_a}{(p_a s_{ac} + (1 - p_a) s_{ah})} > 0,$$

$$\frac{\partial \rho_{Da}}{\partial s_{ah}} = \frac{-(1 - p_a) p_a r_a s_{ac}}{(p_a s_{ac} + (1 - p_a) s_{ah})^2} < 0,$$

$$\frac{\partial \rho_{Da}}{\partial s_{ac}} = \frac{(1-p_a)p_a r_a s_{ah}}{(p_a s_{ac} + (1-p_a)s_{ah})^2} > 0$$

Thus, “improvements” that come about through decreases in p_a and r_a and increases in s_{ah} will lower prevalence. But improvements in s_{ac} will raise prevalence—which is the dark cloud associated with better survival for those with chronic conditions.

We acknowledge that several key concepts are ignored or only present implicitly in the model. For instance, we do not discuss the duration of disease nor the effect of duration on disability and mortality. For notational simplicity, we have also suppressed time. The reader should consider all the model’s parameters and relationships as functions of time period. Indeed, the point of the model is to aid in tracking historical changes in the key determinants of disability decline.

Even though it has limitations, this model captures four important features of the historical record. First, all evidence points to a significant delay in the onset of chronic conditions across the life-cycle—a long-term downward shift of the p_a curve. Evidence we show later indicates that some conditions that were important in the 19th century have been virtually eliminated among those aged 45-64. Delayed onset of chronic conditions is represented as *decreases* over time in p_a at all ages. We do, however, reiterate that increases in industrial pollution and urbanization may have been accompanied by periods of outward shifts of the p_a curve.

Second, the past century has seen well-documented decreases in mortality at all ages. This effect is captured by increases in the s_{ac} and s_{ah} curves. We disagree with Costa’s (2000) conclusion that the life expectancy associated with chronic conditions has remained unchanged since the early 1900s. The duration *conditional on surviving long enough to be examined* may have been high in the 19th century, but many conditions such as heart attacks, strokes, or many

cancers, would have killed far too quickly to allow the individual to enter the Union Army data as having a chronic condition. Still, these people had a condition that would be chronic and debilitating had they survived the acute event. In other words, improvements in treating diseases such as heart attack should be considered as an upward shift in s_{ac} —the mortality schedule of those with debilitating conditions—because the ability to live longer with a heart attack will raise disability prevalence, not lower it.

Our model also implies that even though individuals may have no apparent debilitating condition, prior acute events, in many cases, have lowered their survival such that we can classify them as having a debilitating condition subject to survival curve s_{ac} . Gains in preventing and treating acute infections and accidents increase s_{ah} and, therefore, *lower* disability prevalence. But we need to remember the important contribution of s_{ac} , which *increases* prevalence. The lesson is that changes in disability prevalence are not just a function of the overall mortality schedule, but the conditional mortality schedules discussed above.

The third important historical change concerns r_a , the probability that a condition will be disabling at a given age. Advances in medical technology and the reduced physical demands of work and home life have resulted in large decreases in r_a . But it is possible that these have been partially offset by increasing expectations of health. As population health improves, individuals may be more likely to see a given condition as limiting. And as discussed above, social norms can reinforce this effect, with psychological disorders being the most noteworthy example. A depressed or anxious individual in the 1890s would be seen as a person with character flaws—now we see him/her as disabled.

Fourth, the model illustrates that both the magnitude and slope of the disability age-profile are interconnected functions of non-linear relationships and that age-specific prevalence

need not move monotonically over time. A body of research shows that work-disability was modestly increasing from about the 50s until 1980 (Colves and Blanchet, 1981; Feldman, 1983; Chirikos, 1986). Our model can easily account for this, as well as possible increase in both disease rates and disability during the end of the industrial revolution in the 1890s and early 20th century. The partial derivatives given above also point to the fact that the magnitude of the effect caused by one of the variables is a complex function of the other variables in the model and will not be constant over time.

Figure 2 below illustrates an example of how disability prevalence can transition over periods of time. For simplicity, we define a new term z_a , which is the probability of having a disability at age a : $z_a = p_a r_a$, and we set $s_{ah} = 1$ for convenience. The figure shows the disability prevalence, ρ_a , that would exist for people of age a given values of z_a and s_{ac} . The bold values and the associated arrows shows a hypothetical path that might occur in a population that transitions from a high-mortality, high disease environment to a low-mortality, low disease environment. In this example (as is likely the case of the actually 20th century experience in America), both the disease environment and the mortality schedule are consistently improving (or at least not worsening). This is represented by southwesterly movement across the figure. Obviously a variety of other transition paths could be shown that would tell different stories about disability prevalence that are all associated with a general improvement in the underlying conditions.

		Probability of disability condition ($z_a = p_a * r_a$)								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Probability of being alive, given condition (s_{ac})	0.1	0.011	0.024	0.041	0.063	0.091	0.130	0.189	0.286	0.474
	0.2	0.022	0.048	0.079	0.118	0.167	0.231	0.318	0.444	0.643
	0.3	0.032	0.070	0.114	0.167	0.231	0.310	0.412	0.545	0.730
	0.4	0.043	0.091	0.146	0.211	0.286	0.375	0.483	0.615	0.783
	0.5	0.053	0.111	0.176	0.250	0.333 A	0.429	0.538	0.667	0.818
	0.6	0.063	0.130	0.205	0.286	0.375	0.474	0.583	0.706	0.844
	0.7	0.072	0.149	0.231	0.318	B	0.412	0.512	0.620	0.737
	0.8	0.082	0.167	0.255	0.348	0.444	0.545	0.651	0.762	0.878
	0.9	0.091	0.184	0.278	0.375	0.474	0.574	0.677	0.783	0.890

Figure 2. Age-specific disability prevalence: A hypothetical transition over time.

The example presented in Figure 2 also captures features that seem evident from the historical record. First, there can be sharp declines in disability prevalence as conditions are avoided (or as treatments for conditions improve). Second, periods of apparent increase in disability prevalence can occur even though the general health conditions are improving; this happens when gains in life-expectancy outpace reductions in disease incidence. Third, when survival is high, improving the disease environment (reducing z_a) can result in substantial reduction in disability prevalence, implying that high survival rates are not an impediment to reducing disability in modern populations. Finally, looking at disparate points in time may belie significant changes in the underlying conditions. For example, part of this hypothetical transition involves moving from point A to point B (from a prevalence of .286 to .330 to .250 to .286 again). Comparing A and B implies stability, whereas following the pathway implies cyclicity. But both those interpretations would mask the consistent improvement in both z_a (from .5 to .4) and s_{ac} (from .5 to .6). The over-riding lesson of this example is that a better understanding of disability transitions always accounts for the age-profile and, more importantly,

recognizes that the trend in prevalence is always a function of trends in the underlying components of prevalence.

3. Data

3.1 The Union Army sample

Historical data for this analysis are drawn from the records of the Union Army pension program. The pension system was instituted in 1862 during the Civil War and for the first few decades consisted of men who could convince the Pension Bureau that their disabilities were both severe and related to service in the War. In 1890, a major change in the law was implemented and the program became a general disability program, where disabilities of all etiologies were compensated as long as they reduced the veteran's capacity for manual labor and as long as they were not the result of "vicious habits."

In the years following the 1890 liberalization, the program gradually became a *de facto* old age assistance program and by 1904 became so *de jure*. In 1890, however, the bulk of the Civil War veterans were not old enough to qualify for an age-based pension. Indeed, most were still in their early 50s, and a sizable minority were actually younger than age 50. To qualify for disability compensation, veterans were required to make a formal application to the Pension Bureau and submit to a detailed physical examination by a board of three examining surgeons. These "surgeons' certificates" comprise a vast array of demographic, socioeconomic and, above all, detailed health information on the recruits who applied for pensions. Since veterans could apply for increases in their level of support or take advantage of changes in the laws and procedures governing the system, examinations occurred with some frequency. Veterans had, on average, 4.5 examination records over the course of their lives.

The liberalization of the law in 1890 opened the floodgates for applying for pension assistance. In fact, 50% of those who were ever to be examined for pension assistance were examined in 1891 alone. High rates of examination also existed for neighboring years; about 71 percent of veterans who would at some point enter the pension system were examined between 1890 and 1892. Because the pension system was very generous relative to wages at the time, and because it was not means tested, veterans had strong incentives to apply for assistance.

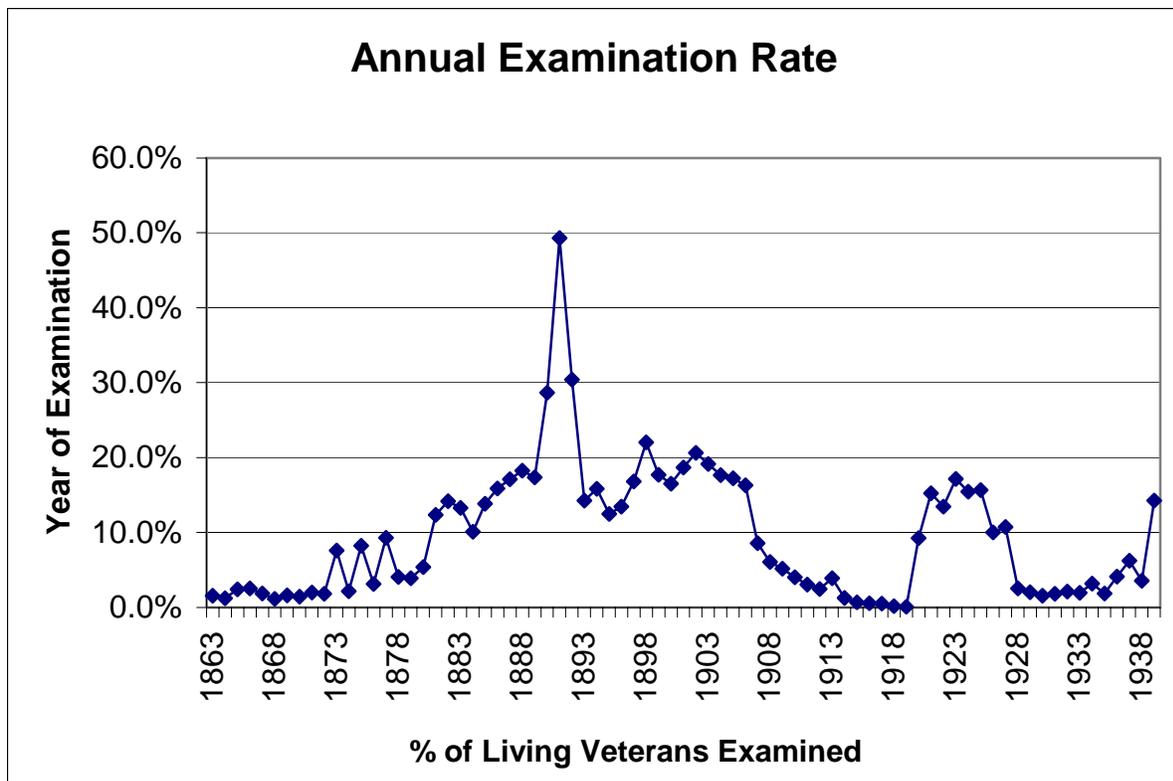


Figure 3.

Thus the sample of exams following the 1890 law constitute a relatively comprehensive cross-section of recruits who believed they had a disabling condition. It is also an ideal time period from our perspective because it is early enough to capture many recruits who are still in their late

40s and early 50s, a period of the life-cycle that is missed when looking at the health data that comes in 1900 or thereafter.

3.2. The NHIS

For the purpose of making comparisons with modern data, we use data from the National Health Interview Survey (NHIS). Each year the NHIS takes a random cross-section of over 100,000 adults in the United States. The NHIS is widely used in research and is the nation's primary source of data for tracking the prevalence of chronic conditions and disabilities in the population. Although the NHIS has been underway since the late 50s, microdata is available from the National Center for Health Statistics only since 1969. We calculate work-disability prevalence at two points in time: 1970 and 2000. Additionally, in order to calculate the prevalence of diseases leading to disability, we combine data from 1997-2001 for our analysis of age-specific prevalence of specific debilitating conditions.

3.3. Disease and Disability Measures

The surgeon's exam provided medical certification of the veterans' health. As part of the exam, physicians rated the applicant's level of disability. This was done either by reporting a total level of disability or, more often, by assigning disability ratings for each condition or combination of conditions. A typical comment would be "Rated \$4 for rheumatism and disease of heart," for example. They were instructed to assign a disability rating only if they felt that the condition in some way limited the veteran's capacity for manual labor. Because of extreme changes in the ratings system over time (as well as conventions in recording disabilities), we create a simple dichotomous variable that indicates whether or not a veteran was given any rating

for disability. We also create dichotomous variables that indicate which conditions the physicians classified as disabling. A variety of information on the health status of veterans is given in the physical exams that have diagnostic value, but our estimates of condition prevalence among the disabled uses only those conditions that the physicians explicitly noted were disabling.

For the Union Army sample, we calculate two different prevalence measures. The first is the prevalence for the entire sample. The second removes from the analysis those cases that are a result of war-disability, so that we can more directly see the impact of the war on disability rates among veterans. Veterans who are disabled in 1893 are classified as war-disabled if they have evidence of injury from their military service records or if they entered the pension system by 1870. Our assessment of the medical exams is that many of the diseases, such as arthritis or heart disease, had a dubious connection to the war, even though physicians routinely made this connection. Medical knowledge at the time was advanced enough to diagnose a variety of chronic illness, but physicians at the time knew almost nothing about the causes of disease; for instance they were not schooled in the germ theory of disease, nor did they understand the important role of genetics, smoking, or environmental toxins (of course, many mysteries about causation remain until this day). Our assessment of war-disability strikes a middle ground between ignoring the unique effects of Civil War combat and taking at face value the physicians' assessment of war-related etiology.

The NHIS underwent a major redesign in 1982 and again in 1997. Definitions of disability and question wording changed in many respects between the two surveys. However, the 1970 and 2000 questions regarding work disability can be compared directly. The 1970 questionnaire asked people about limitations in their "major activity" and in other activities, but

the questionnaire was structure in such a way that men were always asked about their capacity to perform paid work. In 2000, all respondents were asked if they were limited in any way regarding the type or amount of work they could do. Individuals who reported a work limitation were then asked what health condition was the cause of their limitation. The disability measure we use from the NHIS is a dichotomous indicator of whether or not the individual experiences any work limitation.

4. Methods

4.1. Accounting for peculiarities of the Union Army data

In order to make legitimate comparisons to modern data, we have to account for peculiar features of the Union Army database. One of the most important of these has not been highlighted in the previously literature using the Union Army data. For reasons not yet fully uncovered, the disability rating system experienced a massive disruption in 1893. For a period of about 18 months, the practice of noting disease-specific ratings on the examination certificates virtually disappeared, which means that we can exploit physician measurement of disability only up until early in 1893. Because of this data limitation, we report prevalence estimates as of January 1, 1893. Although our estimates are not strictly point prevalence proportions (everyone was not examined on the same day), a large majority of veterans had been examined since 1890, and it is a safe assumption that, because nineteenth century medicine provided few effective treatments of chronic illness, disabilities can be treated as permanent.

Another category of problems has to do with veterans who are missing from the sample as of 1893. Missing cases can be sub-divided into three groups. The first group are those who were not in the system by 1893 but who entered at a later date. Because the pension system was

liberalized in 1890, because the pension awards were substantial in dollar terms, and because there was widespread knowledge about the availability of the pension program, we assume that individuals not in the system were non-disabled. This is a conservative assumption that has the net effect of understating the true prevalence in the cohort.

A second group of missing cases are those who never entered the system, and so we know nothing directly about whether they were alive or what their disability status may have been in 1893. It is important, however, that we do not simply drop these cases from the analysis, since that would overstate the true prevalence of disability. We assume, for the same reasons as given above, that if they were alive they were non-disabled. We account for their possibility of being dead by weighting each case by their probability of being alive as estimated from historical life tables (Haines, 1998) based on their age and the last recorded date in their records (such as a discharge date from the military). Thus if the life table indicates an 80% chance that the veteran is alive, he contributes 80% of a person to the number of non-disabled persons and to the number of persons in the populations. These life table adjusted numbers are then used in the calculation of prevalence rates, and the sample sizes reported in the tables are, consequently, estimated sample sizes.

The third group of missing cases are those who have been examined in the past, but for whom there is no date of death. For instance, a veteran examined in 1891 would be in the data set, but if he has a missing death date, it is not known whether he was still alive in 1893. For these cases, we assume that their disability status is the same as it was at their last exam, but we adjust their contribution to the prevalence ratio by the probability of their still being alive in 1893.

Finally, two problems present themselves with respect to age. First, age determination in the 19th century is always a challenge. To our knowledge, our study is the first to use all the evidence—a possible 81 indicators of age from military service records, census records in 1850, 1860, 1900 and 1910, pension application records, and physical exam records—to determine the best available age for the veteran at a point in time.⁶ During the war, teenagers who wanted to serve had the incentive to overstate their age. In later life, those same recruits had the incentive to overstate their age to increase their probability of pension awards. And those working with census and other 19th century records know the varieties of errors (such as heaping around round numbers) that occur with a high frequency. Second, the age distribution of the Union Army sample in 1893 is a direct function of the age distribution of the Union Army. Recruits were primarily in their late teens and early 20s during the war, which lasted from 1861 to 1865. Furthermore, the age structure of the population in 1893 was significantly different than in 2000. In our analysis, all results are standardized to the US population in 2000 as reported in the 2000 census of the population.

4.2. Potential biases

We argue that most of the unknown biases associated with comparing these two samples will actually serve to understate the Union Army prevalence rates relative to the NHIS. We discuss several of these issues below.

⁶ Because the birth dates were rigorously verified by the pension bureau and were again verified when the dataset was compiled, the age based on the birth date was taken to be the most accurate indicator of each recruit's age. However, in many cases the birth date is non-existent, or is vastly different from the other indicators of age. In fact, over 60 percent of recruits in the Union Army files have a discrepancy of over 1.5 years across the various indicators of their age. About one-third of the recruits have a discrepancy of more than two years. We analyzed all the ages available from each recruit to establish what age among all the indicators of their age was most accurate. We measured the central tendency among their ages, using the median age when the birth date was non-existent, or when the birth date was an extreme outlier, given the other indicators of age present for any given recruit.

Probably the main difference between the two samples is that the Union Army records we use are physician-assessments of disability, defined as the capacity to perform manual labor, and the NHIS measures only self-reported work-disability, with no physician certification. Furthermore, unless a veteran actually showed up to be examined, he is treated as having no disability (weighted by his probability of being alive, as discussed above). Self-reported data is likely to show a higher prevalence because there are some conditions that a physician would not regard as disabling, and individuals may be more likely to rationalize their behavior—people who have withdrawn from the labor force claim they have a work limitation, while workers claim they are not limited (some people with severe conditions, such as paraplegia, will report no work limitations, presumably because they do not like to think of themselves as limited). On balance, we expect that, other things equal, self-reporting increases the prevalence of disability relative to physician reports.

Another significant difference is that the Union Army cohort consists entirely of veterans. Veterans of the Civil War differ from their peers in two important ways: 1) they were all healthy enough in their youth to have become eligible for service; 2) they experienced injury, trauma, and hardship during the war that affected their health over the remainder of their lives. In contrast, in 2000, childhood and adolescent disability is rare, and only a small percentage of those 45-64 have experienced the effects of war. For this reason we calculate the two different prevalence measures, where one excludes war-disabilities.

Another relevant issue is differences in medical knowledge between 1890 and 2000. Looser definitions of disease likely resulted in an over-diagnosis of some conditions. We suspect, for instance, that “rheumatism” and “heart disease” were diagnoses given to people who seemed to be debilitated, but for unknown reasons. On the other hand, coronary artery disease,

which is such a frequent killer in modern populations, could not be diagnosed by the physicians (though they detect valvular disease and cardiac hypertrophy), nor could most cancers or hypertension. For these reasons, the prevalence of particular types of conditions is less reliable in the Union Army data than it is today. But our primary emphasis is on the rate of disability, not the prevalence of chronic conditions. Physicians may have been inaccurate in identifying causes of disability, but they had a variety of indicators to determine whether the person was actually disabled—emaciation, muscle tone, gait, and other signs and symptoms of reduced physical capacity.

Our methodology for extracting prevalence from the Union Army system is likely to understate prevalence because of the time lag between the recruit's last exam and January 1, 1893. If the individual becomes disabled but is not able to be examined before the above date, he is assumed to be healthy. A related problem is that we are treating people who have not entered the pension system as healthy, and we weight their contribution to the prevalence by their probability of being alive based on historical life tables. However, because people with disease have a lower life expectancy than the non-diseased, we are overweighting healthy people in our analysis. Because disease-specific disability rates are not available, we feel it is more conservative to take the current approach rather than try to adjust for differential mortality risks.

4.3. Other analytical limitations

A better understanding of the disability transition would result if we were to directly estimate the prevalence of specific chronic conditions and then show how those conditions result in disability. A copious amount of diagnostic information is present in the medical records to undertake such an exercise. However, this approach is not wise with the Union Army data for the cases we are looking at—namely, younger men in the early 1890s. Men under 65 with non-

debilitating conditions would not be eligible for the pension by 1893 and, therefore, would be less likely to apply and be examined. We would also like to determine accurately the age of onset of both chronic conditions and disability, since this would provide important information relative to two key aspects of the disability transition—an increase in average age of onset of several conditions and a reduction in the debilitating effects of disease. However, average age at onset cannot be determined because people were not eligible for the pension until 1890 if their disability was not related to the war.

Additionally, the differential mortality for people with and without chronic conditions would help us decompose the determinants of disability decline over the twentieth century. We are constrained in our mortality analysis for two reasons. First, we can look at mortality only for those who die after 1890, which limits our understanding of the mortality regime that existed for the decades prior to 1890 and which created the conditions relevant to disability prevalence in 1893. Second, because many important diagnoses are going to be missing from the Union Army data, as noted above, we cannot reliably estimate the impact of those conditions on mortality. Conditions are either fundamentally unobservable or they kill so quickly that the individuals would not have time to enter into the baseline of a survival analysis.

Finally, our comparisons are limited to white males ages 45-64. Similar comparisons with women and minorities are of great interest. Currently data is being prepared at the Center for Population Economics at the University of Chicago that provides the same kind of detailed information on thousands of black recruits who served in the Union Army. We know of no historical microdata on the health status of women, other than vital statistics such as death certificates, which are not systematically available until well into the 20th century.

5. Results

5.1. Disability prevalence estimates

Table 1 below provides our comparisons of disability prevalence between the Union Army cohort and the NHIS. To limit the impact of service in the Civil War on our estimates, we estimate the Union Army disability rates in two ways. First we estimate disability for all veterans in the sample, and then we re-estimate the results excluding any cases that were a result of war-disability, as discussed earlier. We also limit the NHIS to white males, since the Union Army files do not contain African American companies. By both measures, the disability rates in the Union Army sample are much higher than in 1970 and 2000. If we compare the results averaged across age groups, the Union Army veterans without war-disability had a disability prevalence of about 3.5 times higher than men in 2000. We also note that the age profile of

TABLE 1: Work Disability Prevalence, by Age

<u>Age</u>	Union Army, 1893				NHIS			
	<i>All Cases</i>		<i>Excluding War-Disability</i>		<i>1970</i>		<i>2000</i>	
	<u>N</u>	<u>Prev.</u>	<u>N</u>	<u>Prev.</u>	<u>N</u>	<u>Prev.</u>	<u>N</u>	<u>Prev.</u>
45-49	7,093	50.3%	4,838	41.6%	3,084	10.1%	2636	9.5%
50-54	6,841	52.7%	4,327	42.1%	2,762	13.0%	2358	10.3%
55-59	4,061	55.4%	2,589	45.8%	2,496	18.9%	1814	14.8%
60-64	2,405	56.0%	1,565	48.6%	1,991	27.9%	1430	18.6%
All	20,400	52.8%	13,319	43.4%	10,333	16.4%	8,238	12.5%

Notes: Union Army estimates assume those without surgeon's exams had no disability, but the contribution of those missing values is weighted by their probability of being alive on Jan. 1, 1893, based on historical life tables. The NHIS sample is restricted to white males for comparability to the Union Army sample. Both the Union Army and NHIS sample results are age-standardized to match the 2000 U.S. population

disability became steeper between 1893 and 1970, but flattened considerably by 2000. Another way to describe this change is that the prevalence among those aged 45-49 had stabilized by 1970, but that significant declines still occurred for older age groups between 1970 and 2000. Indeed, for those age 60-64, prevalence fell by 33% (from 27.9% to 18.6%); this downward shift is more remarkable in light of evidence (cited earlier) that work disability for those under 65 was increasing during the 1970s.

5.2. Diseases Associated with Disability

Table 2 provides a detailed comparison of the types of conditions associated with disability. Note that these numbers reflect the prevalence of debilitating conditions *among the disabled*, not in the population as a whole. As noted earlier, the Union Army conditions are specifically identified as disabling by the physician. Actual rates of prevalence of these conditions, therefore, will be higher than the rates that we report. The NHIS values are also restricted to those cases where a work disability is present, and only those conditions that are identified as the cause of disability are included. Thus, the non-debilitating (from the view of the respondent) conditions are not included, and the overall prevalence of these conditions among the disabled are higher.

The most striking difference between the two time periods is the prevalence of gastrointestinal conditions as a cause of disability. For instance, 25% of veterans who received a disability rating were rated for hemorrhoids; 15-17% were rated for diarrhea (which is highly co-morbid with hemorrhoids). Other gastrointestinal diseases were rated as well. These numbers compare to only 2-3% in the modern cohorts. On the other hand, psychological disorders were not even rated in the Union Army data, though today they are seen as a major source of

disability. These stark differences between the surveys reveal the importance of both environmental changes and changing notions of the nature of disability.

Table 2: Prevalence of Debilitating Conditions among Disabled Individuals

	Union Army, 1893				NHIS, 1997-2001			
	45-49	50-54	55-59	60-64	45-49	50-54	55-59	60-64
Vision	6.2%	6.2%	7.9%	9.0%	6.0%	7.0%	6.4%	6.5%
Hearing	4.2%	4.9%	6.2%	8.2%	2.5%	4.7%	4.5%	3.8%
Musc./Arthritis	19.2%	20.2%	22.8%	25.1%	44.2%	43.0%	44.8%	39.1%
Injury	5.5%	4.6%	6.0%	5.6%	18.6%	18.0%	15.7%	14.8%
Cardiovascular	14.5%	15.3%	16.4%	19.2%	12.3%	18.7%	25.2%	30.7%
Stroke					2.4%	3.4%	5.0%	7.4%
Hypertension					5.9%	9.1%	9.5%	10.5%
Varicose Veins	1.4%	1.7%	2.6%	4.4%				
Diabetes					8.5%	10.3%	12.2%	12.0%
Respiratory	12.4%	11.2%	11.2%	10.3%	6.5%	7.6%	12.3%	13.7%
Cancer	0.3%	0.5%	0.3%	0.4%	2.0%	3.3%	5.4%	6.9%
Psychological					16.6%	13.0%	9.2%	5.3%
Nervous	4.3%	4.5%	4.1%	5.3%	7.0%	6.6%	5.5%	4.8%
Gastrointestinal	19.5%	19.7%	20.7%	20.1%	2.3%	3.0%	3.4%	2.0%
Hernia	4.7%	5.6%	6.2%	6.7%				
Genitourinary	2.6%	2.9%	3.4%	4.5%	1.5%	0.9%	1.9%	1.8%
General/Other	8.5%	9.9%	11.0%	14.3%	12.1%	12.5%	10.6%	8.3%
N=	4838	4327	2589	1565	1463	1616	1632	1629

Notes: These figures represent the prevalence of debilitating conditions among individuals who are disabled. They do not include other conditions that the disabled individual may have had. For the Union Army data, conditions include only those for which a specific disability rating was assigned; similarly, the NHIS are conditions identified by the respondent as causing his disability. The Union Army figures are for those where cases of war-disability have been removed (see text).

The levels and age patterns associated with cardiovascular and respiratory disease are consistent with the conjecture we make above about both high incidence and high relative mortality. Cardiovascular disease among the Union Army veterans was relatively constant across the age groups, while it is sharply increasing (from 12.4% to 30.6%) among the NHIS

group. Similarly, the prevalence of respiratory diseases more than doubles from 6.5% to 13.9% among the NHIS sample, but the proportion actually falls from 21.6% to 15.7% in the Union Army sample. Thus the incidence of these conditions is high enough to generate high prevalence levels, but the mortality is also sufficiently high to suppress or even diminish cross-sectional increases by age.

Other conditions typically not thought of as highly fatal are also generally consistent with our conjecture about mortality. We would expect the prevalence of conditions that people could live with for long periods of time to be increasing with age. (though if these conditions are frequently co-morbid with more life-threatening conditions, the age profile would be relatively flat). We note that non-fatal conditions such as hearing problems, vision problems, varicose veins and “general debility,” which consists mostly of musculoskeletal conditions not classified as “rheumatism” (arthritis), are all generally increasing with age.

6. Discussion

6.1. Understanding the 20th century decline in disability

To summarize the results above, we present three basic findings related to work disability. First, we find sharply higher prevalence (roughly 3-6 times as high, depending on age) among Union Army veterans aged 45-64 than exists at the present. Second, the age-profile is surprisingly flat, even flatter than the modern profile, rising from 41.6% for ages 45-49 to 48.6% for ages 60-64. This is flatter than the 2000 profile and only half the slope of the 1970 profile. Third, the distribution of conditions that are the cause of work disability has shifted substantially. A variety of conditions are highly prevalent in the 19th century, but injuries and gastrointestinal disorders are particularly important among the Union Army cohort.

Musculoskeletal problems (mostly arthritis), cardiovascular disease, and psychological disorders are the most important in the modern population.

Our findings of a dramatic long-term decrease in the level of work disability implies that many of the factors that may have increased disability prevalence (such as increases in s_{aC} or r_a) were swamped, over the long term, by factors that decrease prevalence (decreases in r_a or p_a). That disability prevalence has declined so sharply in an era of falling mortality rates is evidence of the tremendous strides that have been made with respect to chronic illness—both over the past century—both in terms of incidence and in mitigating the debilitating effects of disease. We strongly suspect that the probability of a functional limitation resulting in a disability (the F_4 function described earlier) has fallen among working-age adults in the past century, mostly due to the fact that work is much less physically demanding than it was in times past.

An age-profile that is high in magnitude but whose slope is only modestly increasing points to three important features of 19th century work-disability. First, disability was highly prevalent across the life course, not just in older ages. Our results suggest an average age at onset that was much younger than today (average age at onset cannot be calculated directly because we are missing medical exams on most of the cohort prior to the 1890 liberalization of the law). An early onset of debilitating conditions would, naturally, lower the incidence at later ages (though severity might increase substantially with age). The high rate of rejection of Union Army recruits for health reasons is consistent with high rates of adolescent and young adult disability. Fogel (1994) claimed that the rejection rate for recruits examined for service in the Civil War was 30.1%. The wide range of conditions that is affecting even those men as young as 45-49 (as seen in Table 3) indicates early onset of several debilitating conditions.

Second, the types of conditions that led to disability are consistent with a relatively flat age-profile. These are conditions that tend to be non-fatal and that can strike early. For instance, conditions such as hernias and hemorrhoids were routine causes of disability in the 19th century. Conversely, heart disease—so important today among those in the 55-64 year age group—tends to strike later and, more importantly, is associated with high mortality. Cardiovascular disease is a much more important cause of disability among those 55-64 today than it was for the Union Army veterans, most likely because people are surviving so much longer with heart disease today than in the 19th century.

Third, the evidence is consistent with general trends in mortality. Table 3 shows selected period life table values for 1890, 1970 and 2000. For each year, the ${}_nq_x$ schedule gives the probability that a person alive at the beginning of the interval dies by the end of the interval, and the l_x schedule represents the survival curve, the number of people alive at the beginning of the age interval per 100,000 live births. As we have noted earlier, high mortality suppresses disability prevalence, but it also *flattens* the age profile of disability since the slope of s_{dC} with respect to age becomes steeper as mortality increases (it is even possible that prevalence can decline with age). Reductions in mortality at all ages have continued since 1970, which indicates great strides in disease prevention and accommodation since that point.

In terms of the recent trend, the large reductions in disability prevalence among the 55-64 year olds shown in Table 1 are particularly noteworthy, especially since this group also had the greatest reduction in mortality (which, other things equal, would increase disability). This points to dramatic delays in the onset of disability among people in this age group. Some might attribute the decline in work-disability to the dramatic decrease in labor force participation of men in this age group (the labor force participation rate fell from 83.0% in 1970 to 67.7% in

Table 3: Selected Life Table Values: 1890, 1970, and 2000

Age Interval	1890		1970		2000	
	${}_nq_x$	l_x	${}_nq_x$	l_x	${}_nq_x$	l_x
20-25	0.035	73,362	0.010	96,482	0.006	98,604
25-30	0.039	70,799	0.008	95,526	0.006	97,977
30-35	0.045	68,045	0.009	94,721	0.007	97,363
35-40	0.054	65,047	0.013	93,847	0.010	96,675
40-45	0.057	61,771	0.021	92,632	0.014	95,755
45-50	0.069	58,237	0.034	90,705	0.021	94,441
50-55	0.085	54,237	0.054	87,647	0.031	92,474
55-60	0.116	49,616	0.085	82,952	0.048	89,731
60-65	0.154	43,851	0.127	75,886	0.078	85,586
65-70	0.216	37,086	0.184	66,231	0.124	79,419
70-75	0.294	29,091	0.255	54,029	0.200	70,657
75-80	0.414	20,526	0.357	40,258	0.322	58,874

The ${}_nq_x$ curve represents age specific mortality, the probability that a person alive at the beginning of the age interval will die before the end of the interval. The l_x curve is the survival curve, the number of people still alive at the beginning of the age interval from 100,000 live births. Sources: 1893: Haines (1998); 1970: NCHS (1972); 2000: Official NCHS life table (Aries, 2002)

1990, though it has been stable since 1990). However, the NHIS in 2000 asks respondents whether they are limited in the type or amount of work they can do—whether or not they are working. Researchers such as Waidman, Bound, and Schoenbaum (1994) have argued that people who have left the labor force would be *more* likely to rationalize their decision by reporting themselves as having a work-disability, especially if they are receiving SSDI (and SSDI participation has *risen* in 1970). Instead, we find significant declines in the rate of work-disability in this age cohort.

The model of the disability age-profile we have described is consistent with a variety of alternative scenarios of how the position and shape of the disability age profile may have changed over the past century. As noted earlier, work-disability was likely increasing from the

1950s until about 1980. Periods of increase such as this are entirely consistent with the model we present and occur because increases in s_{aC} were greater than the increase in $p_{a}r_a$. The shape of the profile was also considerably steeper in 1970 than in either 1893 or 2000, implying that the earliest gains in disability came first at younger ages (45-54). Since 1970, the gains have been more concentrated among the older ages (55-64). The net effect of these changes suggests a transition illustrated in Figure 4 below.

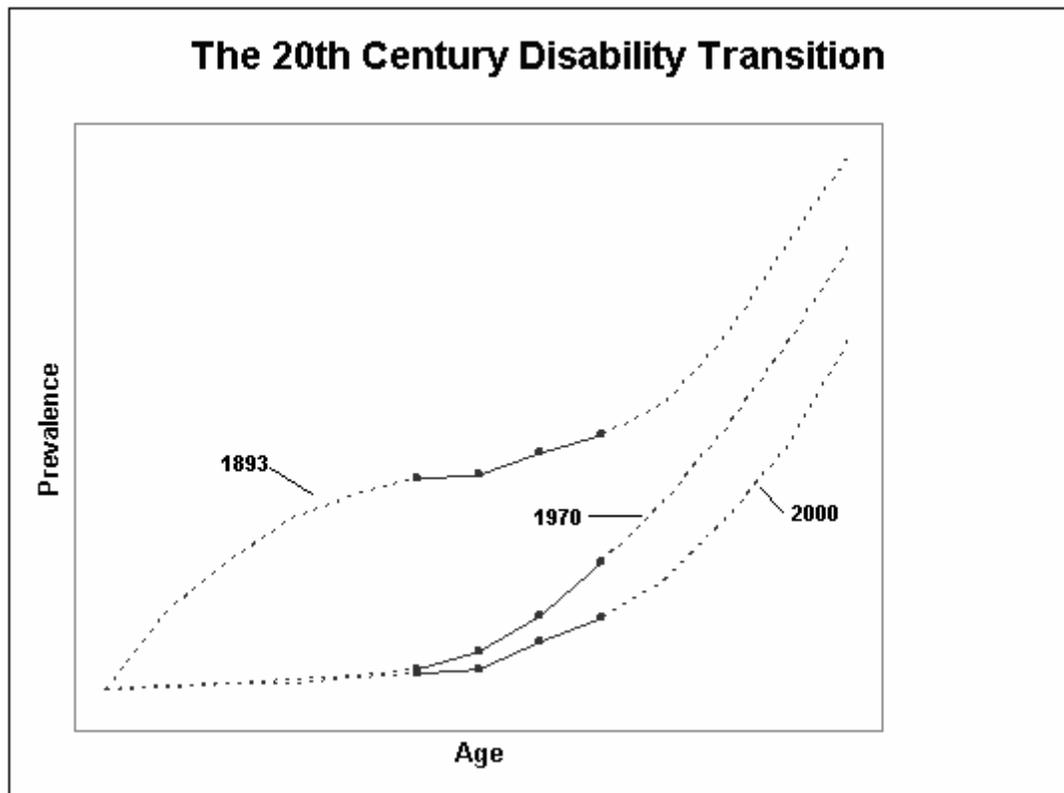


Figure 4. The 20th Century Disability Transition

The solid portions of the curves above show our results for ages 45-64 of the last section, while the dotted lines are only conjectures. In this graph, we are deliberately vague about assigning disability levels to particular ages, which is why we do not have a scale for age and prevalence on the axes. But the dashed lines are consistent with the high rate of younger age disability in the historical period and the decline in disability at older ages found in numerous studies cited

earlier. We also re-emphasize the point that increases in prevalence may have occurred for particular ages during the period of time covered.

In summary, Figure 4 shows the long-term transition that has occurred in work disability over the past century. The greatest gains were early in the century, but the decline was also significant over the last three decades as well. Indeed, disability is becoming steadily less important as a health concern for working age men and, we suspect, for women as well. The concern that increasing life expectancy would worsen health is clearly not relevant for the population under age 65 unless there are forces that would reverse that trend. We take up this topic next below.

5.2. Implications for future trends

Even though there were some documented increases in the prevalence of work-disability in the 1970s, these trends have been reversed and levels of work-disability are substantially lower in 2000 than in 1970. But can the work-disability of non-elderly adult males continue to fall? In short, we think there are compelling reasons to anticipate further declines, though certain underlying trends point to potential increases. Of course trends for the ages we have looked at in this paper may not move in tandem—the prevalence may move in different directions for different age groups.

Perhaps the most important reason to expect further declines in work-disability is that the mortality rates, as shown in Table 3, have very little room for improvement, especially in the 45-54 year-old range. Substantial declines in mortality across the life cycle have been a powerful force pushing *upwards* on disability, but this force has been masked by gains in the prevention and accommodation of chronic illness. To illustrate how mortality is becoming less of force on

disability prevalence, consider the following thought experiment—supposed that age-specific mortality rates were to be cut in half at every age from their 2000 levels (a change that would raise life expectancy at birth for white males by a whopping 8 years) and that the “new” people resulting from this mortality reduction would have twice the prevalence of disability as their 2000 counterparts. How would these relatively dramatic events change the prevalence of work-disability? Our calculations show that disability would rise by .3 percentage points at age 45-49, by .5 at 50-54, by 1.0 at 55-59 and by 1.8 at 60-64. Only among the oldest age group 60-64 could we possibly expect reductions in mortality alone to increase disability prevalence in any meaningful way—but this is the group for which the declines in work disability since the 1970s have been most pronounced! Furthermore, while reductions in mortality among those with chronic conditions push disability upward, reductions in mortality due to acute conditions and accidents will continue to push prevalence downward. In short, the mortality story in relation to work-disability is largely over.

Because mortality has little role left to play in disability prevalence prior to age 65, the question becomes whether we can expect further improvement on the disease front. The answer here depends on which specific conditions we are talking about. Reductions in the prevalence of vision and hearing problems, stroke, cancer, neurological disorders, and genitourinary disease would have minimal effects on the overall disability prevalence, especially for the 45-54 year-olds because, individually, these conditions have minimal effect on disability prevalence. On the other hand, better treatment of arthritis could have significant effects, where it is a contributing cause of disability for almost half of the disabled persons aged 45-54. Continuing to push off the incidence of heart disease (which is a cause of disability for 12.5% of 45-49 year-olds and 30.5% of 60-64 year-olds) could also cause significant reductions in disability for those in the older age

groups. Clearly there are several potential advances in both the prevention and treatment of disease that could further suppress disability among working age adults.

A key determinant of disease incidence continues to be behavior. The model presented earlier was not a behavioral model, but we would expect health to be treated as a normal good, so that increasing real income should reduce disability, holding relative prices fixed. However, recent decades of rising real incomes have seen sharp increases in the rate of obesity.

Lakdawalla and Philipson (2002) argue that technological changes have increased obesity through making home and market production more sedentary and through lower relative food prices.

In any case, the two major causes of work-disability are arthritis and heart disease, both diseases for which obesity is a significant risk factor. Because the obesity increases have occurred among adolescents and younger workers, in addition to older workers, future generations of workers in the 45-64 age group may face significantly higher rates of work-disability than their predecessors. However, the silver lining to Lakdawalla's and Philipson's results is that the same increase in sedentary work activities that underlie the obesity increases are also likely to accommodate less healthier workers, thereby lowering the rate of work-disability. In other words, society may be producing a generation of overweight adults with high rates of disease and functional limitations, but it is also producing a workplace where physical capacities are increasingly less important.

Even though smoking has been declining (another reason for increasing obesity (Rashad and Grossman, 2004)), the prevalence of asthma has risen significantly in the last decade. Lakdawalla, Bhattacharya and Goldman (2001) analyze the increase in disability as measured by the ability to perform personal care. They find an increase in personal care disability among

people in their 40s from 1984 to 1996, though the prevalence of this relatively severe type of disability was still under 4% among people in their forties in 1996. They claim that the increase in asthma prevalence among this younger group is sufficient to explain the entire increase in this type of disability. The lesson is that future trends are a function of trends in specific debilitating conditions. Increases in asthma, diabetes and obesity among the young should remain as important potential determinants of future disability trends. Furthermore, it may be the case that health declines among younger workers may have potent effects as these individuals age, as opposed to the types of conditions whose incidence occurs at later ages.

Finally, disability prevalence is intrinsically linked to the interaction of an individual's demand for health and his or her labor supply. While some of the earlier withdrawal from the labor force in recent decades has been due to more generous social insurance benefits, this decline is largely a function of non-health factors, such as higher quality and cheaper leisure activities (Costa, 1998). The decline in labor force participation and the declines in the physical demands of labor have both been important determinants in the decline in work-disability. The steep declines in male labor force participation in the 70s and 80s have stabilized since 1990, and there are no reasons we know of to expect an imminent departure from stability. The likeliest change in the future will probably have less to do with labor supply decisions than they will be caused by technological innovations that continue to reduce the physical demands placed on human bodies in the workplace.

7. Conclusions

The epidemiological transition, in which acute infectious diseases were largely eliminated as a cause of death among young people, led scholars to worry that the rise in chronic

illnesses as causes of death would be associated with high rates of disablement. As the title of Verbrugge's influential 1984 paper queried, will we have "longer life but worsening health?" The long-term answer to this question is a resounding "no." The last century has seen dramatic declines in the prevalence of work-disability across all ages from 45-64. The hypothesis of worsening health is consistent with the sharp increases in the survival curve over the past century, but associated with the mortality decline have been improvements in the prevention and treatment of disease. Our evidence shows that many chronic conditions that were a routine cause of disability in the 19th century have been largely eliminated as causes of disability in the year 2000. Furthermore, diseases when they occur are less likely to become disabling because of both treatment and rehabilitative care and because of changes in the workplace that have reduced the physical demands of labor. While we look only at men ages 45-64, abundant published evidence indicates that disability among older adults has decline substantially over the past century as well. At some point disability prevalence may increase at very late ages, but it is too early to tell when that might occur or at what ages.

Even though work-disability prevalence has declined sharply over the century, including in the time period since 1970, we find compelling reasons to believe that further gains will be made among working-age men. Technological innovation is likely to provide advances in both medical care and in improving workplace conditions, which will both further suppress disability, as well reductions in mortality from acute conditions and accidents. Moreover, the unseen but important role that mortality from chronic illness has played in pushing disability upwards has virtually run its course, especially among those 45-54. Therefore, future trends in work-disability will be a function of disease and its consequences. Gains in preventing and treating arthritis and cardiovascular disease will determine, in large part, the future declines in work-

disability. However, even though we have become used to continuing medical advances in our day, the troubling increase in obesity among children and adults poses a serious risk, since obesity is a risk factor for several debilitating diseases.

Most recent research has neglected the issue of work disability in favor of studying trends in ADL disability, especially among the elderly. But we argue that in the aggregate, the most important type of disability for younger adults is probably not ADL disability but, instead, the pain, discomfort and limitation of a variety of activities—work being particularly important—usually undertaken by ones peers. It is clear that for a large number of working-age males in the 19th century, pain and disability were common aspects of life. An important topic for additional research is to understand long-term trends in disability among women and minorities, which we were unable to study with the Union Army data. Large scale reductions in disability should be seen as contributions made by medical and public health innovations (as well as improved working conditions and reduced hours of work) that, perhaps, rival any improvements in the standard of living measured by indices such as GDP. Moreover, these dramatic changes are likely to continue to shape important social and policy variables such as retirement decisions and the utilization of health care for years to come.

References

- Arias E. (2002). United States life tables, 2000. *National vital statistics reports; vol 51 no 3*. Hyattsville, Maryland: National Center for Health Statistics.
- Chirikos, T.N. (1986). Accounting for the historical rise in work-disability prevalence. *Milbank Quarterly*, 64, 271-301.
- Colvez, A. & Blanchet (1981). Disability trends in the United States population 1966-76: Analysis of reported causes. *American Journal of Public Health*, 71, 464-471.
- Cost, D.L. (1998). *The Evolution of Retirement: An American Economic History 1880-1990*. University of Chicago Press for NBER.

----- (2000). Understanding the twentieth century decline in chronic conditions among older men. *Demography*, 37, 53-72.

----- (2002). Changing chronic disease rates and long-term declines in functional limitation among older men. *Demography*, 39, 119-138.

Crimmins, E.M. (1996). Mixed trends in population health among older adults. *Journals of Gerontology*, B51, S223-S225.

Crimmins, E.M. & Saito, Y. (2000). Change in the prevalence of diseases among older Americans: 1984-1994. *Demographic Research*, 3, Article 9.

----- (2001). Trends in healthy life expectancy in the United States, 1970-1990: gender racial and educational differences. *Social Science and Medicine*, 52, 1629-1641.

Crimmins, E.M., Saito, Y. & Reynolds, S.L. (1997). Further evidence on recent trends in the prevalence and incidence of disability among older Americans from two courses: the LSOA and the NHIS. *Journal of Gerontology*, 52B, S59-S71.

Cutler, D.M. (2001). The reduction of in disability among the elderly. *Proceedings of the National Academy of Science*, 94, 2593-2598.

Cutler, D.M. & Richardson, E. (1997). Measuring the health of the United States population. *Brookings Papers on Economic Activity, Microeconomics*: 217-71.

Feldman, J. (1983). Work ability of the aged under conditions of improving mortality. *Milbank Memorial Fund Quarterly* 61:430-444).

Fogel, R.W. (1994). *Early Indicators of Later Work Levels, Disease, and Death*. Grant proposal submitted to National Institutes of Health by the Center for Population Economics, University of Chicago. p. 264.

Fogel, R.W., & Costa, D.L. (1997). A theory of technophysio evolution, with some implications for forecasting population, health care costs, and pension costs. *Demography*, 34, 49-66.

Freedman, V.A, Martin L.G, & Schoeni R.F. (2002). Recent trends in disability and functioning among older adults in the United States: A systematic review. *Journal of the American Medical Association*, 288, 3137-3146.

Freedman, V.A. & Martin, L.G. (1998). Understanding trends in functional limitations among older Americans. *American Journal of Public Health*, 88, 1457-62.

----- (2000). The contribution of chronic conditions to aggregate changes in old-age functioning. *American Journal of Public Health*, 90, 1755-61.

- Fries, J.F. (1980). Aging, natural death and the compression of morbidity. *New England Journal of Medicine*, 303, 130-135.
- Gruenberg, E.M. (1977). The failure of success. *Milbank Quarterly*, 55, 3-34.
- Haines, M. (1998). "Estimated Life Tables for the United States, 1850-1900." *Historical Methods*. 31(4): 149-69
- Lakdawalla, D, Bhattacharya, J. & Goldman, D. (2001). Are the young becoming more disabled? *NBER Working Paper 8247*.
- Lakdawalla, D. & Philipson, T. (2002). University of Chicago, George J. Stigler Center for the Study of the Economy and the State Working Paper no. 174.
- Manton, K.G. & Gu, X. (2001). Changes in the prevalence of chronic disability in the United States black and non-black population above age 65 from 1982 to 1999. *Proceedings of the National Academy of Science USA*, 98, 6354-6359.
- Manton, K., Corder, L. S., & Stallard, E. (1997a), Chronic disability trends in elderly United States populations. 1982-1994. *Proceedings of the National Academy of Sciences of the USA: Medical Sciences*, 94, 2593-2598.
- Nagi, S.Z., (1965). Some conceptual issues in disability and rehabilitation. In Sussman, M.B., ed., *Sociology and Rehabilitation*. Washington, DC: American Sociological Association.
- National Center for Health Statistics (1972). *Vital Statistics of the United States, 1970: Volume II, Section 5, Life Tables*. Rockville, MD.
- Ostchega, Y. Harris, T, Hirsch, R., Parsons, V.L., Kington, R. (2000). Prevalence of functional limitations and disability in older persons in the US: data from the National health and Nutritional Examination survey. *Journal of the American Geriatric Society*, 48, 1132-1135.
- Pope, A.M. & Tarlov, A.R. (1991). *Disability in American: Toward a National Agenda for Prevention*. Washington, DC: National Academy Press.
- Rashad, I & Grossman, M. (2004). The Economics of Obesity. *The Public Interest*, Issue #156 National Affairs, Inc.
- U.S. Bureau of the Census (1975). *Historical Statistics of the United States, Colonial Times to 1970, Bicentennial Edition, Part I*. Washington, DC.
- U.S. Bureau of the Census (2003). *Statistical Abstract of the United States*. Washington, D.C.
- Verbrugge, L.M. (1984). Longer life but worsening health? Trends in health and mortality of middle aged and older persons. *Milbank Quarterly*, 62, 475-519.

Verbrugge, M. and Jette, A. M. (1994). The disablement process. *Social Science and Medicine*, 38, 1-14.

Verbrugge, L. M., Gates, D.M., & Ike, R.W. (1991). Risk factors for disability among US adults with arthritis. *Journal of Clinical Epidemiology*, 44 (2), 167-182.

Waidman, T., Bound, J. & Schoenbaum, M. (1995). The illusion of failure: Trends in the self-reported health of the US elderly. *Milbank Quarterly*, 73, 253-287.

Wilson, S.E. (2003). The Prevalence of Chronic Respiratory Disease in the Industrial Era: The United States, 1895-1910, “ in Dora Costa (ed.) *Health and Labor Force Participation over the Life Cycle: Evidence from the Past*, Chicago: NBER and University of Chicago Press.