Survival and the Changing Role of Chronic Diseases in the 20th Century[1]

Frank Sloan, Duke University and NBER, Martin Salm, University of Mannheim, and Dan Belsky, Duke University

Substantial improvements in health and longevity occurred during the 20th Century in the U.S. and in other industrialized countries. These improvements in health have plausibly led to substantial improvements in economic-well-being not reflected in national income accounts. Nordhaus (2005) has argued that between 1950 and 2000, improvements in health have been as valuable as all other sources of economic growth combined.

The resulting increase in longevity has led to an older population which is at higher risk for chronic disease and associated disability. Given the reduced threat of infectious disease by the late 20th Century, these chronic diseases and functional limitations were more likely to be causes of death. However, some severely chronically ill persons who would have been kept alive in 2000 plausibly died more quickly in the late 19th Century, leaving a lower proportion of severely ill and disabled persons in the surviving population. This is shown rather dramatically in this study comparing determinants of longevity with 1900 and 1992 as baseline years, using longitudinal data for each period from Union Army pension records and the Health and Retirement Study (HRS).

[1] Partial financial support for this research came from the National Institute on Aging grant 2R37-AG-17473-05A1.

I. Data

We use data used from two sources. The first is longitudinal data based on records from the Union Army pension program.[2] Prior to enactment of Social Security, this was the largest pension program in the U.S., covering 85% of Union Army veterans by 1900, and 90% by 1910 (Costa 1998, p. 160). The program began in 1862, providing pensions to volunteer and regular recruits severely disabled as a direct result of military service. In 1890, the program was broadened to encompass all disability. Although old age was not recognized by statute as a sufficient qualification for receiving a pension until 1907, a minimum pension was granted to all those aged 65+ before this year unless they were "unusually vigorous" (Costa 1995). The mortality experience of this sample of veterans is representative of the white male population as a whole in 1900, indicating that mortality at older ages for these veterans was largely a product of environmental factors, rather than wartime experience (Costa et al. 1998, pp. 197-212).

The "Surgeon's Certificates" data (ICSPSR Series 3417) contain medical records used by the Bureau of Pensions to evaluate pension applications, and include results of physical examinations performed by three physicians. These records are matched with the 1900 Census of Population, from which we obtain information on date of birth, and marital and occupational status and place of residence. Entry into the veterans sample was contingent on examination by the panel of physicians. Persons were retained in the sample irrespective of whether they received a pension. The total Surgeon's Certificate sample includes 17,280 white male individuals who applied for a pension by 1900. 3,931 individuals died before 1900. An additional 1,303 observations are dropped, mostly because these men were under age 16 in 1865 or lacked information on dates of birth and

[2] See Costa (2000) for more detail.

death. Our analysis sample consists of 11,622 white males aged 51+. Surgeon's Certificate data on the file were grouped into 21 primarily organ system-based health screens. Our analysis uses entry screen/recruit information, cardiovascular, endocrine disease, general appearance, nervous system, respiratory, and rheumatism/musculoskeletal screens. Data on dates of death come from Union Army pension records. All individuals for whom we have data, including those for whom a link to the 1900 U.S. Census could not be established, are included in our analysis.

Although the Union Army data extend from the 1860s through the 1930s, we use 1900 as the baseline year in our analysis of veterans' data. Having a few survivors into the 1930s provides a good follow-up period for the analysis of mortality. Because of the change in policy at the pension board, many more individuals chose to be examined after 1890. By 1900, this information had likely spread throughout the veteran community, yielding a more representative sample of the mature males than if an earlier year had been selected as the baseline year. A veteran does not enter the sample until he applies for pension benefits. In addition, 1900 is the year for which we have occupation and marital status information on these individuals from the U.S. Census.

For comparison purposes, we use data from the first seven waves of the Health and Retirement Study, 1992-2004 with 1992 as the baseline year. Our analysis sample consists of 3,595 white males aged 51+ at baseline.

II. Specifications of Hazards

We estimate mortality hazards, assuming a Weibull distribution and allowing for unobserved heterogeneity.[3] The hazard function at time *t* for individual i with covariate vector X_{it} is,

$$\lambda(t \mid X_{it}, \theta) = \lambda_0(t) * \exp(X'_{it}\, \theta) * \eta_i, \qquad (1)$$

where $\lambda_0(t) = \alpha * t^{\alpha-1}$ is the baseline hazard for the Weibull distribution with the shape parameter α, and $\exp(X'_{it}\, \theta)$ is the proportional hazard with parameters θ. We account for individual-specific unobserved heterogeneity by including a multiplicative term η_i (>0), distributed as gamma with mean 1 and variance σ (see e.g., Lancaster 1979).[4] Failure is defined as death of the individual. We know the year of the death from pension records for the veterans. For HRS, for which data are only available in two-year intervals, we know the person's age at baseline; deaths are assumed to have occurred in the year between adjacent interviews. All survivors to the 2004 interview, as well as individuals who were lost from the sample but who could not be verified as dead (N= 813) are treated as censored.[5] To make the veterans' analysis more comparable to that for HRS, we estimate a hazard using all data on deaths, and also a hazard based on exclusively on data for 1900-1912.

We include the following explanatory variables in the Union Army analysis: congestive heart disease (oedema,cyanosis,dyspnoea); other heart disease (valvular heart disease (aortic or mitral murmurs), and arteriosclerosis (or atherosclerosis), chronic respiratory disease (emphysema, bronchitis, pneumonia, tuberculosis, bronchiectasis,

---

[3] Lancaster (1990) provides an excellent review of the estimation of duration models.

[4] Three assumptions required to account for unobserved heterogeneity are: (1) the heterogeneity term is independent of the covariates; (2) the heterogeneity term has a distribution known up to some parameters; and (3) the heterogeneity term enters the hazard function multiplicatively. The mean also needs to be normalized to one for identification. We adjust the standard errors by clustering by individuals.

[5] The individuals lost to the survey are treated as censored one year after the last interview.

fibrosis, hemoptysis, pleuritis); stroke (cause of nervous disorder is stroke or aphasia); cancer, diabetes (any symptom of diabetes from physician's comments or from urinalysis), arthritis (musculoskeletal problem with mention of joint or ligament or tendon), pain in bladder/ kidney, mental illness (anxiety, delusion, depression, histrionic, hypochondria, paranoia, passive-aggressive, schizophrenia, suicidal, homicidal, mania, antisocial, violent, hysteria, phobia), low cognition (delirium, dementia, mental retardation, poor memory, senility), limitation walking (any entry about gait), requires physical assistance due to general debility; body mass index (BMI), referring to information from the last examination before 1900: obese, BMI>30, overweight, 25<BMI < 30, underweight, BMI< 20, normal, omitted); married (1900 U.S. Census), occupation (following Wilcox 1992), with professionals/managers, the omitted category (1900 U.S. Census); and size of city of residence (>250,000, 26,235-250,000, omitted category <26,235 in 1909 (U.S. Census Bureau 1909) . A health condition is noted if it was present on any examination record before 1900 except for BMI for which we take the most recent value before 1900.

We attempt to make the specification for the HRS analysis as comparable to the specification for the veterans as possible. Differences in specifications result from differences in available data. Fortunately, the results are reasonably robust to changes in specification.

For the analysis with HRS we include the following explanatory variables, all self-reported by respondents as of the baseline in 1992: congestive heart failure, other heart problem (heart attack, coronary heart disease, angina, other); chronic respiratory disease (lung disease such as chronic bronchitis, emphysema); stroke; cancer (except skin cancer); diabetes; arthritis; kidney/bladder problems; mental illness (emotion, nervous, psychiatric problems); low cognition (lowest quartile of scores on HRS cognition test); limitation walking (walking across room at least “somewhat difficult”); body mass index (same as in veterans analysis); married; occupation (manual service, artisan, professional (either based on occupational designation or with a college degree, with managers the omitted group); lives on farm or ranch; and education (highest degree college+, some college, high school, omitted group < high school). Educational attainment is used in an alternative specification which drops the binary variables for occupation.

III. Descriptive Characteristics

The white male Union Army veterans were 61 years of age in 1900 on average versus 56 years of age for white men in the HRS sample in 1992 (Table 1). Several chronic diseases were more prevalent among men in this age group in 1900 than in 1992: congestive heart failure; other heart disease; chronic respiratory disease; and stroke. Plausible reasons for the differences vary by disease. About half of the veterans’ sample had other heart disease, largely valvular heart disease, attributable to infectious diseases owing to lack of availability of anti-infectives.

Rates of cancer, diabetes, arthritis, mental illness, and low cognition were far lower in 1900. However, cancers could not be diagnosed unless the tumor was visible. Thus, many cancers were undoubtedly missed. By 1900, physicians diagnosed diabetes from urinalysis. Between 20% and 50% of whites received urinalysis at medical examination in 1895 and between 60% and 80% in 1900 (Humphries et al. forthcoming).

Rates of obesity were far lower in 1900 than in 1992. Twenty percent (19.9%) of white males were obese in 1992 compared with 3.7% in 1900 (adjusting the Table 1 value for the fraction of missing values for BMI). Mental illness was recorded in only 1.1% of cases; it may have been underdiagnosed, one reason being a greater stigma associated with this condition. Moreover, all but the most severe forms of depression were unlikely to have been diagnosed as mental illness in 1900. The much lower rate of low cognition in 1900 than in 1992 plausibly reflects differences in the way this is measured in the two data bases.

Being limited in walking was much more prevalent in 1900 than in 1992, but the fractions of men receiving personal assistance in daily activities of daily living is about the same in both years. Farming was an important occupation in 1900. Because an implausibly small number of persons in the HRS reported farming as an occupation (N=20), we use those who reported living on a farm or ranch (N=201, 5.5%) instead.

IV. Estimated Mortality Hazards: 1900 and 1992

Mortality hazard regressions are similar using either a 12-year follow-up for the HRS or a follow-up until death in the veterans' analysis (Table 2). The 12-year follow-up is presented to permit a comparison with the 12-year follow-up for HRS.

The hazard ratios for the chronic diseases tend to be much higher in the HRS than in the veterans' analysis, the exception being for mental illness for which they are very similar. For other heart, the hazard ratios are similar.

Persons with severe chronic disease for which there was no effective therapy in the late 19th Century probably died before 1900, leaving those with less severe chronic disease alive. Open-heart surgery for persons with severe valvular disease, the only cure for this condition, was first introduced in the late 1960s. Also there was a dramatic shift in the predominant type of heart disease from valvular to ischemic during the 20th

Century. Dialysis for end-stage renal disease was first introduced in the mid 1940s and only became routinely available to the public with the inclusion of dialysis as a Medicare benefit in 1972.

Limitations in walking were much more common in 1900 than in 1992; however, the effect on survival was much smaller in the veterans' than in the HRS analysis. Possibly some veterans successfully feigned having difficulty in walking to get a pension in 1900. Increased BMI substantially decreased longevity in 1900. Obese men were about 40% more likely to die in a year than were normal weight men. By contrast, being obese did not affect longevity in the HRS analysis. But a much higher percentage of males were obese in 1992 than in 1900. Thus, being obese made one much more of an outlier in the earlier year. Being married increased longevity in both cohorts.

In both years, being "unproductive" (not employed, the term used in the Union Army data), decreased longevity, more so in HRS than in the veterans' analysis. In the veterans' analysis, depending on the specification, manual and service workers and artisans tended to have higher hazards of dying than did professionals or managers, the omitted reference group. Occupation did not influence longevity in the HRS analysis, although men living on farms and ranches had appreciably lower mortality. Among veterans, being a farmer was about equivalent to being a professional or manager in terms of longevity; living in a very large city decreased survival.

Although occupation does not affect longevity in the HRS analysis, educational attainment has a positive effect on longevity. The 1900 U.S. Census did not collect data on educational attainment.

V. Survival Probabilities

Survival probabilities were lower at the beginning than at the end of the 20th Century (Fig. 1). The death rate per 100,000 individuals was halved between these two years.

The Kaplan Meier curves show nonparametric estimates of survival probabilities for 51 year old white males up to age 75 for the HRS sample, and up to age 90 for the Union Army sample. Few men in the 1992 HRS sample are observed beyond age 75. The probability of an average 51 year old Union Army veteran surviving to age 75 is 42.4%, compared with 78.7% in the HRS sample. The reduction in mortality is less pronounced for men with chronic diseases. With a (non congestive) heart condition at age 51, survival probabilities to age 75 are 41.6% for Union Army veterans and 72.0% for HRS respondents. For those with a chronic respiratory disease, the corresponding probabilities are 31.7% for the veterans and 41.3% in the HRS.

The effect of chronic heart and respiratory conditions on longevity is much larger in the analysis for the later compared to the earlier period. However, the types of heart diseases and lung diseases which were most common in 1900, such as valvular heart diseases and tuberculosis, are different from today's most common heart and lung diseases. Thus, there is reason for caution in interpreting these findings.

VI. Estimated Values of Eliminating Specific Diseases: 1900 and 1992

We compute potential gains in life expectancy from the elimination of congestive heart disease, other heart disease, chronic respiratory disease, diabetes, limitations in walking, and obesity that could have been achieved in 1900 and 1992 We also calculate gains in life expectancy from marriage.

To calculate potential gains in life-expectancy, we compute annual mortality hazard rates from ages 50 to 99 for respondents with and without one of the above conditions, such that (1) the ratio of the mortality rates for men with and without the condition are equal to the estimated hazard ratio presented in Table 2, and (2) mean mortality rates for men with and without the condition are equal to life table mortality rates. For the HRS analysis, we use life tables for white males in 1998 (www.cdc.gov/nchs/nvsr48_18.pdf, accessed 06/08/05). For the veterans, we apply life tables for white males in 1900 (Haines 1994).

Based on the annual mortality rates we calculate life expectancy at age 50 for men with and without the condition, and for population means. Potential gains to an individual from eliminating a given condition are calculated as the difference between their life expectancy with the condition and their life expectancy without the condition. The potential gains for the total male population are calculated as the difference between average life expectancy and life expectancy without the given condition.

In 1900, the health conditions with the largest impact on life expectancy for men with the condition were obesity--2.7 years, congestive heart disease--1.1 years, and chronic respiratory diseases--0.9 years. In 1992, the health conditions with the largest impact on life expectancy for persons with the condition were diabetes--15.9 years, limitations in walking--8.1 years, and chronic respiratory diseases--6.4 years (Table 3). For the total male population, considering those men with and without these conditions, health conditions with the largest impact on life expectancy were other heart diseases--0.14 years, obesity--0.11 years, and chronic respiratory diseases--0.108 years. In 1992, the health conditions with the largest impact on life expectancy for the total population are diabetes--0.59 years, chronic respiratory diseases--0.38 years, and limitations in walking--0.16 years. Overall, potential gains from eliminating chronic diseases were much larger in 1992 than in 1900, which mostly reflects the decline in competing mortality risk factors over the last century. The gain in life expectancy from marriage is 1.3 years in 1900 and 2.4 years in 1992. If all men had the life expectancy of married men, the life expectancy for the total population at age 50 would have increased by 0.9 years in 1900 and by 1.9 years in 1992.

To calculate the value of increased life expectancy, we assume the value of a year of life to be $100,000 in 1992, both in 1900 and 1992. We also assume that individuals discount the value of future life by a real rate of 3% per year. Based on annual mortality rates we calculated the value of future life for population means and for people with health conditions. We identify potential benefits of increased longevity by taking the difference between the value of life for the population average and the value of life for a person with the health condition. The benefit from eliminating a health condition varies with age. Therefore, we calculate the benefit separately for men at ages 50, 60, 70, and 80.

The total benefit of increased longevity was calculated based on the benefit for men with the health condition at different ages, the total male U.S, population in the age groups 50-59, 60-69, 70-79, and 80+ and the share of men in each age group affected by the health condition. We assume that the shares of men affected by health conditions were constant across age groups, equal to the shares in our baseline samples in 1900 and 1992. Numbers for total U.S. male population by age groups for 1900 and 1990 are from Hobbs and Stoops (2002). All dollar values are expressed in 1992 dollars.

In 1900, the most expensive diseases with respect to the effect on life expectancy were chronic respiratory diseases--$25.7 bil., and other heart diseases--$20.4 bil. In 1992, the most expensive diseases were diabetes--$813.6 bil. and chronic respiratory diseases--$565.7 bil.

VII. Discussion

Survival in late life improves from the Union Army dataset to HRS. Prevalence of some categories of chronic disease and for functional limitations declined from the Union Army sample to that in the HRS. For certain chronic conditions such as diabetes, prevalence increased. In addition to declining prevalence of some chronic diseases, longevity for those with the conditions increased over the 20th Century. The mortality hazards of these chronic conditions, however, increased.

Overall improved longevity in the modern period is attributable to a combination of a variety of factors including better nutrition (Fogel 1997), improved public health and related reductions in infectious disease (Cutler et al. 2006), and better medical care (Cutler 2004, Cutler et al. 2006). That some diseases became more prevalent while others less so may reflect certain of these trends. Cancer, diabetes, and arthritis all increase in prevalence with age (Merrill et al. 2000, Winer and Sowers 2004, Reginster 2002). As a result, an older population will inevitably see more of these conditions. Further, elderly individuals must be reasonably healthy in terms of other disease burdens, for the effects of these conditions to be salient to an examiner. Consequently, as the sorts of infectious diseases that produced chronic respiratory and pulmonary conditions at the turn of the 20th Century declined in incidence, age-related conditions became more apparent. Additionally, the broad categories of disease we define likely include different mixes of illness in the two samples. “Other heart” primarily represents valvular heart disease in the Union Army sample. That same variable is likely composed mainly of ischemic heart disease in the HRS. A similar shift from infectious to age-related conditions can be seen for chronic respiratory illness.

The increased longevity of individuals with chronic diseases in the modern period is likely explained by the tremendous improvements in medical care and technology in the second half of the 20th Century (Cutler et al. 2006). Perhaps less apparent at first glance, is the seeming contradiction between this and our finding that hazards for these chronic conditions are greater now than 100 years ago.

Three principal factors have influenced the increase in hazard of chronic illness over the 20th Century. First, competing risks declined. The increase in mortality due to accidents notwithstanding (with few motor vehicles on the road in 1900), mortality from sources other than chronic illness has declined dramatically. Chief among these is infectious disease, responsible for 34% of deaths in the death registration states in 1900 and only 5% in the U.S. in 1992 (U.S. Centers for Disease Control 2006).

Second, due to increased longevity of those with chronic illness, the very sick remain in the population for longer in the modern period. By contrast, in the sample of veterans, very ill individuals died relatively quickly, most likely failing to survive to 1900. So there is a greater proportion of the "illest ill" in our HRS sample relative to our Union Army sample, resulting in greater hazards due to illness.

Finally, enhanced medical knowledge and improved diagnostic techniques have increasingly enabled physicians to identify conditions such as heart disease far earlier than was possible at the end of the 19th and beginning of the 20th Century. As a result, individuals dying suddenly of heart attacks or even those dying slowly of chronic respiratory diseases may not have been consistently identified as having these conditions in the sample of veterans, resulting in lower hazards. Conversely, higher hazards are obtained from the HRS data in part because almost all of the individuals who have a given illness are diagnosed with it.

We acknowledge these limitations. First are the limited diagnostic techniques available to the examining surgeons in 1900. While the HRS data on diseases come from respondent self-report and the veterans' data on the same comes from physician diagnosis, it seems like that the diagnoses in the veterans' data are less precise than those in the HRS. Second, the definitions of the various diseases have been matched between the two data bases as well as possible, but the matching remains quite imperfect for reasons beyond our control. Third, certain variables such as educational attainment and income are lacking for the veterans' analysis. Fourth, a few results are puzzling, such as why obesity would decrease longevity at the beginning but not at the end of the 20th Century.

References

Costa DL. (1995). “Pensions and Retirement: Evidence from Union Army Veterans.” *Quarterly Journal of Economics*, 110(2): 279-319.

Costa DL. (1998). *The Evolution of Retirement: An Economic History, 1880-1990*. Chicago: University of Chicago Press.

Costa DL (2000). “Understanding Mid-life and Older Age Mortality Declines: Evidence from Union Army Veterans.” National Bureau of Economic Research, Working Paper No. 8000, November.

Cutler DM. (2004). *Your Money or Your Life*, Oxford: Oxford University Press.

Cutler DM, Deaton AS, Lleras-Muney A. (2006). “The Determinants of Mortality.” *The Journal of Economic Perspectives*, 20(3):97-120.

Fogel R. (1997). “New Findings on Secular Trends in Nutrition and Mortality: Some Implications for Population Theory.” Chapter 3 in MR Rosenzweig and O Stark eds. *Handbook of Population and Family Economics*, Cambridge: Elsevier Science.

Haines, Michael R. (1994), “Estimated Life tables for the United States, 1850 – 1900, National Bureau of Economic Research. Historical Working Paper No. 59.

Lancaster, T. (1979). “Econometric Methods for the Duration of Unemployment,” *Econometrica,* 47(4): 939-956.

Lancaster, T. (1990). *The Econometric Analysis of Transition Data*. New York: Cambridge University Press, 1990.

Hobbs, F. and N. Stoops (2002). “Demograpic Trends in the 20th Century.” U.S. Census Bureau, Census 2000 Special Reports.

Humphreys M, Costanzo P, Haynie KL, Ostbye T, Boly I, Belsky D, Sloan F. “Racial Disparities in Diabetes a Century Ago: Evidence from the Pension Files of US Civil War Veterans.” *Forthcoming Social Science and Medicine*.

Merrill RM, Capocaccia R, Feuer EJ, Mariotto A. (2000). “Cancer Prevalence Estimates Based on Tumor Registry Data in the Surveillance, Epidemiology and End Results (SEER) Program.” *International Journal of Epidemiology, 29*:197-207.

Nordhaus WD. (2005) Irving Fisher and the Contribution of Improved Longevity to Living Standards. *The American Journal of Economics and Sociology, 64(1)*:367-392.

Reginster JY. (2002). The Prevalence and Burden of Arthritis. *Rheumatology, 41(suppl. 1)*:3-6.

Wilcox NW. (1992) A Note on the Occupational Distribution of the Urban United States in 1860, in Robert W. Fogel et al. (eds.), *Without Consent or Contract: The Rise and Fall of American Slavery, vol. 2 Evidence and Methods.* New York: W.W. Norton and Co.

Winer N, Sowers JR. (2004). Epidemiology of Diabetes. *Journal of Clinical Pharmacology, 44*:397-405.

U.S. Census Bureau (1900). *Census of Population*.

U.S. Census Bureau (1909). *General Statistics of Cities*. Report for the U.S. Department of Commerce.

U.S. Centers for Disease Control. "Leading Causes of Death, 1900-1998." http://www.cdc.gov/nchs/data/dvs/lead1900_98.pdf. Accessed 06/07/2006.

Table 1: Sample Means

	Union Army Sample	HRS Sample
Age at entry	61.0	55.9
Congestive heart failure	0.049	0.011
Other heart failure	0.496	0.078
Chronic respiratory disease	0.120	0.065
Stroke	0.047	0.020
Cancer	0.002	0.077
Diabetes	0.025	0.067
Arthritis	0.209	0.408
Bladder/kidney problems	0.102	0.121
Mental health problem	0.011	0.091
Low cognition	0.021	0.127
Cognition missing		0.066
Limitation walking	0.282	0.018
Requires assistance	0.024	0.034
BMI < 20	0.141	0.067
25 < BMI ≤ 30	0.173	0.326
BMI > 30	0.036	0.199
BMI missing	0.073	
Married	0.694	0.750
Marital status missing	0.167	
Manual work	0.076	0.042
Service	0.039	0.315
Artisan	0.085	0.029
Farmer	0.228	
Farm hand	0.022	
Unproductive	0.027	0.390
Occupation missing	0.409	0.002
College degree		0.141
Some college		0.206
High school degree		0.405
City pop > 250,000 in 1909	0.037	
Next 139 Cities in 1909	0.034	
Location missing	0.649	
Lives on farm / ranch		0.055
N	11,622	3,595

Table 2: Mortality Hazard Regression†

	Union Army Sample		HRS Sample	
	12 years follow-up (1)	Until death (2)	12 years follow-up (3)	12 years follow-up (4)
Congestive heart	1.217**	1.138**	1.589	1.653
	(0.079)	(0.049)	(0.413)	(0.429)
Other heart	1.058	1.034	1.048	1.052
	(0.033)	(0.020)	(0.182)	(0.182)
Chronic respiratory	1.158**	1.108**	2.230**	2.227**
	(0.052)	(0.031)	(0.333)	(0.329)
Stroke	0.959	0.964	2.704**	2.712**
	(0.070)	(0.044)	(0.579)	(0.580)
Cancer	1.034	1.193	1.581**	1.604**
	(0.329)	(0.250)	(0.252)	(0.256)
Diabetes	0.841	0.996	3.171**	3.155**
	(0.084)	(0.059)	(0.474)	(0.471)
Arthritis	0.969	1.010	0.939	0.932
	(0.036)	(0.023)	(0.109)	(0.108)
Bladder/kidney	1.047	0.996	1.305	1.305
	(0.051)	(0.031)	(0.182)	(0.183)
Mental	1.123	1.037	0.999	1.016
	(0.157)	(0.092)	(0.163)	(0.166)
Low cognition	1.087	1.033	1.530**	1.422*
	(0.107)	(0.070)	(0.206)	(0.198)
Cognition missing			1.247	1.163
			(0.251)	(0.237)
Limitation walking	1.080*	1.039	2.671**	2.999**
	(0.036)	(0.022)	(0.682)	(0.762)
Requires assistance	1.127	1.059	0.948	0.975
	(0.100)	(0.064)	(0.227)	(0.234)
BMI < 20	1.103*	1.043	2.121**	2.226**
	(0.047)	(0.028)	(0.368)	(0.385)
25 < BMI ≤ 30	1.111*	1.133**	0.854	0.835
	(0.046)	(0.029)	(0.118)	(0.115)
BMI > 30	1.430**	1.367**	0.877	0.871
	(0.109)	(0.069)	(0.134)	(0.133)
BMI missing	1.003	1.003		
	(0.061)	(0.061)		
Married	0.817**	0.860**	0.694**	0.730**
	(0.034)	(0.023)	(0.081)	(0.084)
Marital status	1.036	0.989		
Missing	(0.058)	(0.037)		
Manual work	1.109	1.133**	1.227	
	(0.083)	(0.050)	(0.381)	
Service	1.200*	1.078	1.187	
	(0.108)	(0.058)	(0.218)	
Artisan	1.156*	1.060	1.126	
	(0.082)	(0.045)	(0.416)	
Farmer	0.949	0.976		
	(0.060)	(0.035)		

Farm hand	1.050	1.039		
	(0.121)	(0.071)		
"Unproductive"	1.194	1.056	1.588**	
	(0.111)	(0.066)	(0.277)	
Occupation missing	1.069	0.999	2.452	
	(0.063)	(0.034)	(2.484)	
College degree				0.562**
				(0.119)
Some college				0.682*
				(0.119)
High school degree				0.795
				(0.104)
City pop > 250,000	1.194*	1.132*		
in 1909	(0.063)	(0.058)		
Next 139 Cities	1.015	1.058		
in 1909	(0.090)	0.056		
Location missing	0.972	1.024		
	(0.023)	(0.039)		
Lives on farm /			0.557	0.547
Ranch			(0.189)	(0.186)
N	11,622	11,622	3,595	3,595

† Weibull hazard model, with gamma distributed heterogeneity term; standard errors in brackets
* significant at 5% level, ** significant at 1% level

Table 3: Potential Gains in Life Expectancy From Elimination of Health Conditions*

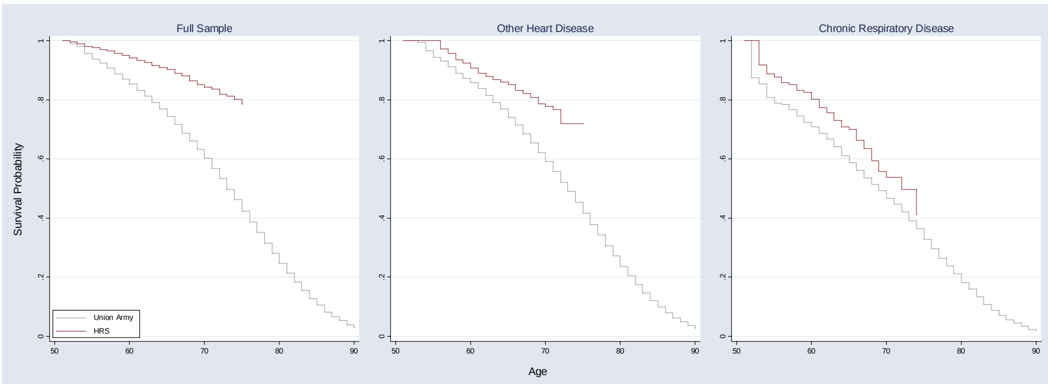
	Men with Condition		Total Male Population	
	Union Army	HRS	Union Army	HRS
Congestive heart failure	1.105	3.981	0.057	0.040
Other heart disease	0.281	0.277	0.140	0.021
Chronic respiratory disease	0.873	6.361	0.108	0.380
Diabetes	0.000	15.586	0.000	0.591
Limitation walking	0.322	8.098	0.092	0.159
Obese	2.687	-0.639	0.110	-0.143
Married**	1.278	2.374	0.405	0.488

* All estimates are for white men at age 50
**Gain if all persons were married

Table 4: Value of Potential Gains in Life Expectancy from Eliminating Health Conditions (in Billions of 1992 $)[+]

	1900	1992
Congestive heart	14.3	65.6
Other Heart	20.4	40.2
Chronic respiratory	25.7	565.7
Diabetes[++]	0.0	813.7

[+] We assumed the value of a life-year to be 100,000 in 1992 USD for both periods, and an individual discount factor of 3%. The economic benefits were calculated based on the gains in life expectancy for persons with health conditions, the share in the population affected by health conditions, and the size of the male U.S. population in the relevant age groups.
No effect on life expectancy
[++] For 1900, the value of eliminating diabetes is zero, because our estimated effect of diabetes on life expectancy is zero.



F0igure 1: Kaplan-Meier Survival Estimates