1999

The DES Analysis/Evaluation Plan: Task 4, Report 8

H. Allan Hunt
W.E. Upjohn Institute, hunt@upjohn.org

Kelly DeRango
W.E. Upjohn Institute

Richard V. Burkhauser
Syracuse University

J. S. Butler
Vanderbuilt University

Westat, Inc.

Task 4: Report 8

The DES Analysis/Evaluation Plan

Submitted to:

Social Security Administration
Office of Acquisition and Grants
1710 Gwynn Oak Avenue
Baltimore, Maryland 21207

By:

W.E. Upjohn Institute for Employment Research
300 South Westnedge Avenue
Kalamazoo, MI 49007-4686

and

Westat, Inc.
1650 Research Boulevard
Rockville, MD 20850

August 20, 1999

Table of Contents

List of Tables

List of Figures

I. *Introduction to Analysis/Evaluation Plan*

This section introduces Westat's approach to meeting the policy research objectives for the main study, as distinguished from the predominately methodological objectives for the pilot study. The main study objectives are broken down further into four core objectives and four objectives related to the redesign of the disability determination process. A brief overview of the proposed methods for addressing each objective is offered in this introduction together with a tabular presentation that summarizes our approach to all objectives. In the next section, each objective will be discussed in more detail. The final section describes our approach to meeting the redesign objectives. It should be understood that this report represents a description of our plans at this very early point in the study. Many modifications and adaptations will doubtless be necessary as data collection proceeds and more experience accumulates.

A. Overview of Study

The Disability Evaluation Study (DES) was solicited under SSA-RFP-98-3102. The RFP specified the proposed work plan and the research design questions to which the Social Security Administration (SSA) sought answers.

The four **core objectives** for the DES are:

1. Providing an estimate of the U.S. population of persons potentially eligible to receive disability benefits;
2. Determining how many persons in the U.S. population do not meet eligibility requirements today, but may meet them at some point in the future;
3. Identifying factors that enable disabled individuals to remain in the work force (e.g., treatments, accommodations, and interventions);
4. Identifying self-reported measures appropriate to estimating future program eligibility.

In addition to these broad research objectives, SSA enumerated a set of objectives that would support and evaluate their ongoing disability determination process redesign efforts.

The following **redesign objectives** will also be analyzed as part of the DES:

1. Testing and validating the redesigned decisionmaking process as implemented by SSA;

2. Estimating the impact of the redesigned decisionmaking process on program size and costs;

3. Assessing changes in the number and kinds of persons who might be found disabled under the redesigned process, relative to the current process;

4. Identifying self-reported measures appropriate to estimating future program eligibility under the redesigned decisionmaking process.

A random sample of 100,000 persons will be screened and sub-sampled according to four strata: (A) 500 SSA beneficiaries, (B) 3,000 seriously impaired nonbeneficiaries, (C) 1,500 borderline impaired nonbeneficiaries, and (D) 500 nonimpaired nonbeneficiaries. Data collected from approximately 10,000 respondents to the DES comprehensive survey instrument (CSI) and 5,500 respondents who complete physical exams will provide the information base with which these objectives will be addressed by the Westat team. Completion of the survey instruments plus the physical exams, mental assessments, performance tests, and functional assessments according to the DES protocol will provide the richest database yet available to address critical disability policy issues.

Matching DES data to SSA administrative data on earnings and program participation will provide additional information for estimating the current and future extent of disability in the United States. The data gathered from respondents in the DES will enable an accurate estimate of the number of persons currently eligible for SSDI and SSI on the basis of their physical and mental condition, i.e., ignoring their current economic status (**Core Objective 1**). This will be done through applying the current SSA/DDS disability determination process to simulated application folders using the data developed in the DES. Westat has also proposed the development of a "gold standard" of disability determination for the DES. The "gold standard" will consist of the best available information on the individual respondent reviewed by the best available expert to determine who is disabled and who is not. Developing and utilizing an achievable "gold standard" of disability assessment will further improve our understanding of who is disabled and why.

In all, the DES will utilize four different measures of disability: (1) self-report of disability status by respondents, similar to that which has been used in other large surveys; (2) simulation of the existing SSA/DDS disability determination process, including appeal to ALJ level; (3) the "gold standard," which will be designed to represent the current best feasible practice under conditions of the DES; and (4) simulation of the "redesigned" disability determination process, as specified by SSA.

The gold standard will also be useful to evaluate self-report measures which will be included in the DES for determining disability (**Core Objective 4 and Redesign Objective 4**) and, possibly, to validate any new decisionmaking process that SSA adopts (**Redesign Objective 1).**[4] Self-reported disability status can be compared to status determined under the "gold standard" to determine how accurate such self-reports are, and what biases may exist in self-reported disability data. By asking the same, or similar, questions as existing surveys for a portion of the self-report, it will also be possible to determine the likely biases in existing data sources. Benchmarking DES measures of disability to such continuing longitudinal surveys also offers a way to introduce a dynamic element into our estimates of future applicants and beneficiaries.

Sophisticated modeling approaches will be used to predict the number of SSDI and SSI applications in the future (**Core Objective 2**). Applying state-of-the-art econometric methods to estimate models of claimant application behavior and disability determination system outcomes based on cross-sectional comparisons should provide an acceptable range of estimates for future claimant population and system costs. Our models of claimant behavior will predict how many individuals will apply to SSDI and SSI programs in each year, given a set of basic determinants. Entirely separate models of the disability determination process will provide estimates of how many of these applicants will actually become beneficiaries and their ultimate benefit levels. Applying these two sets of models to projected values for determinants at future points in time will enable estimates of future SSDI and SSI program size and cost levels.

These same models can also be used to predict the effects of any changes to the current disability determination system (**Redesign Objectives 2 and 3**). We will gather informed opinions ("assumptions") as to the likely impact of specific changes in the disability determination process from individuals familiar with both systems. These "assumptions" will then be used as inputs to the applicant models and the disability determination process models. The models will yield predictions of the number of persons applying, the administrative outcome, and the projections of the number of beneficiaries and program costs.

The information gathered from DES respondents will include questions about access to health care, rehab services, community supports, and employer accommodations that have enabled them to remain in the workforce, return to work, or otherwise remain active despite serious impairments. Comparing DES survey respondents who are working with those who are not, holding physical or mental

---

[4] This assumes that SSA implementation decisions are made before the survey plans are finalized.

impairments equal, will assist in identifying those factors that enable disabled persons to remain in the workforce (**Core Objective 3)**.

A conceptual model of work disability has been constructed to guide development of the instrumentation and analysis of the DES (see figure 1). The model assumes work disability results from an interaction of individual factors and environmental factors that jointly explain beneficiary and employment status. The adoption of a contextually dependent model of work disability is based on currently prevailing research and theory. Factors of the environment (e.g., physical barriers, attitudes, health care, policy) and of the person (e.g., age, education, skills, work values, and resources) combine with medical and functional factors to determine how a given health condition will affect activity and restrict participation in social roles including work (e.g., Scotch & Shriner, 1997; Linton, 1998; Verbrugge & Jette, 1994).

In our conceptual model the individual factors include health status, functional effects of health conditions on daily activities and work, and personal characteristics including work history and education. Environmental factors are comprised of social support, access to community, work environment, access to health care, education and vocational services. All these factors will be assessed in the comprehensive survey instrument that will be administered to some 10,000 individuals representing a range of disability experience across the entire population. Dimensions of human variation and environmental response will be considered together with the characteristics of potentially disabling health conditions to account for the work and benefit outcome categories to which the persons belong.

While we are aware of the manifold influences which help to shape an individual's experience in the workplace, it is not yet possible to incorporate all of these in the same empirical model. Thus, our analysis and evaluation plan seeks to draw its inspiration from the model in figure 1 and stretch the empirical boundaries, while using proven analytical techniques. During the course of the project we will seek ways to integrate more of these factors into our empirical models of disability and work. Table 1 provides a summary of the entire analysis and evaluation plan. It can be used as an overview of the research objectives and our planned approach to each of them. Full details are presented in the text.

Table 1. Summary of Analysis Plan

Core Objectives			
Research Objective	**General Approach**	**Steps**	**Methodology**
1. Estimate US population potentially eligible to receive disability benefits.	Direct simulation of disability determination process.	Simulation: 1. Collect DES data. 2. Develop simulation procedures. 3. Develop a gold standard process. 4. Using simulation procedures, obtain initial DDS decisions. 5. Model ALJ decisions on simulated claim folders (alternate: simulate ALJ decisions using ALJs). 6. Determine interrater reliability of simulation. 7. Determine validity of simulated decisions. 8. Estimate number in US population found disabled in the simulation. 9. Compare simulated decisions and decisions made with gold standard to assess validity.	To determine interrater reliability, make repeat determinations of a subset of simulated claim folders. To assess validity, apply the gold standard to data from sample of respondents.

Research Objective	General Approach	Steps	Methodology
2. Determine number of persons in US population who do not meet eligibility requirements today, but may meet them at some point in the future.	Econometric modeling to indicate the timing of application for disability benefits.	1. Model who will apply to SSI and SSDI. 2. Model who will become beneficiaries. 3. Model the level of benefits for each beneficiary. 4. Apply models (1) through (3) to data from DES respondents. 5. Determine number of beneficiaries and costs of program.	Use reduced form modeling and structural modeling. Project future trends using data from other databases (e.g., CPS, SIPP, Occupational Outlook Handbook, etc.) on demographics, economic conditions, eligibility, generosity of other income-support programs, and earnings from SSA administrative files.
3. Identify factors that enable persons with severe impairments to remain in the workforce.	Compare those who are eligible for benefits (as identified in the simulation) and working with those eligible and not working. Compare those who are severely impaired with those who are moderately impaired. Compare work patterns of beneficiaries before and after increase in SGA.	1. Identify and operationalize dependent variable (working/ not working). 2. Identify and operationalize independent variables (type and severity of impairments, program variables, etc.). 3. Identify and operationalize intervening variables [age, sex, geographic region, occupation, personal economic factors, accommodation, medical interventions, rehabilitative efforts, etc.]. 4. Run models using multiple logistic regression, CART, and other procedures.	Multivariate modeling using multiple logistic regression, CART, and other procedures.

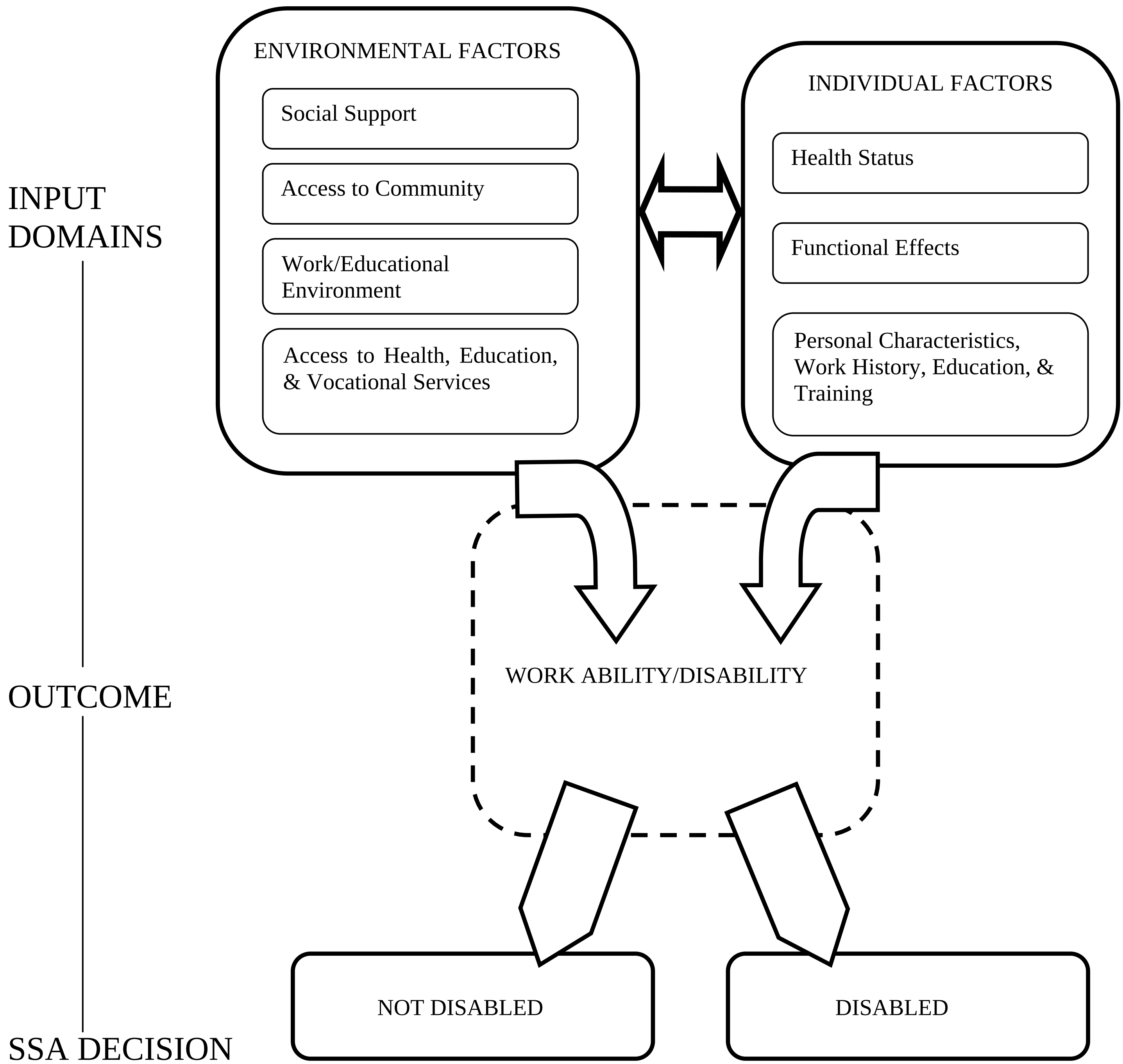
Research Objective	General Approach	Steps	Methodology
4. Identify self-reported measures appropriate to estimating future program eligibility.	Compare decisions modeled with varying combinations of self-report data to decisions made with gold standard.	1. Develop several models based on different types of information collected (e.g., self-reported information from screener and medical questionnaire, results of performance tests. 2. For each model, the dependent variable is disability status (based on results of the simulation); independent variables are different combinations of information; intervening variables are age, race, sex, region or state, industry, occupation, macroeconomic variables and other factors affecting job prospects. 3. Determine the probability of disability using each model. Those with a probability of 0.50 or above (or some other cut-off) will be considered disabled. 4. Compare the decision results (as ascertained by the probabilities) of each participant with the decision results using the gold standard to calculate sensitivity and specificity of each model.	Multiple logistic regression, average derivative estimation.

Redesign Objectives			
Research Objective	**General Approach**	**Steps**	**Methodology**
1. Test and validate redesigned decision-making process as implemented by SSA.	Compare decisions made with redesigned process to simulated decisions made under existing DDS process.	1. Test instruments for use in field. 2. Test content validity of instruments. 3. Develop methods for simulation. 4. Test interrater reliability of simulation process. 5. Test concurrent validity of the process.	Compare simulated decisions made under revised procedures with the original simulation.
2. Estimate impact of redesigned decision making process on program size and costs.	Use models of program application, disability determination, and program cost developed under basic objectives.	1. Model who will apply to SSI and SSDI. 2. Model how many will become beneficiaries. 3. Estimate costs under alternative program.	Use microsimulation methods to model behavioral responses to changes in the decision-making process or other program characteristics.
3. Assess changes in the number and kinds of persons who would be found disabled under the new process, relative to the current process.	Compare results of the redesign simulation with results of the original process simulation.	1. Compare prevalence of disability ascertained in current and redesigned system. 2. Calculate (a) percent found disabled in both systems; (b) percent found disabled in the redesign but not in the current system; and (c) percent found disabled in the current system but not the redesigned system. 3. Compare characteristics of the mismatches (e.g., age, gender, type of impairment, level of functioning, severity of impairment, etc.).	Simple comparisons of beneficiary characteristics according to classification under the alternative processes.

Research Objective	General Approach	Steps	Methodology
4. Identify self-reported measures appropriate to estimating future program eligibility under redesigned decisionmaking process.	Compare decisions modeled with self-report data to decisions made with gold standard.	1. Develop estimation models for predicting eligibility from self-reported data. 2. Determine the sensitivity and specificity of predictive models. 3. Identify ongoing surveys which provide similar self-reported data to enable tracing of future trends in the eligible population.	Multiple logistic regression.

Figure 1

WORK ABILITY/DISABILITY MODEL
(Single Point in Time)



B. Analysis of Current Beneficiaries

INTRODUCTION

The DES is being designed to study the disability programs of the Social Security Administration. But the Social Security Administration administers two separate disability insurance programs – Social Security Disability Insurance (DI) and Supplemental Security Income (SSI) for the disabled (and the aged, which will be ignored for now). These programs have evident differences in the number of awards given over the past 15 years as seen in Figure 2. Awards for SSI recipients have been more volatile than awards for SSDI recipients especially in the 1990s. The Supreme Court decision in *Sullivan v. Zebley* followed by the recession of 1991 saw a tremendous increase in new SSI recipients, while the economic expansion beginning in 1992 corresponded with a decline in SSI awards.

Figure 2. Number of Awards

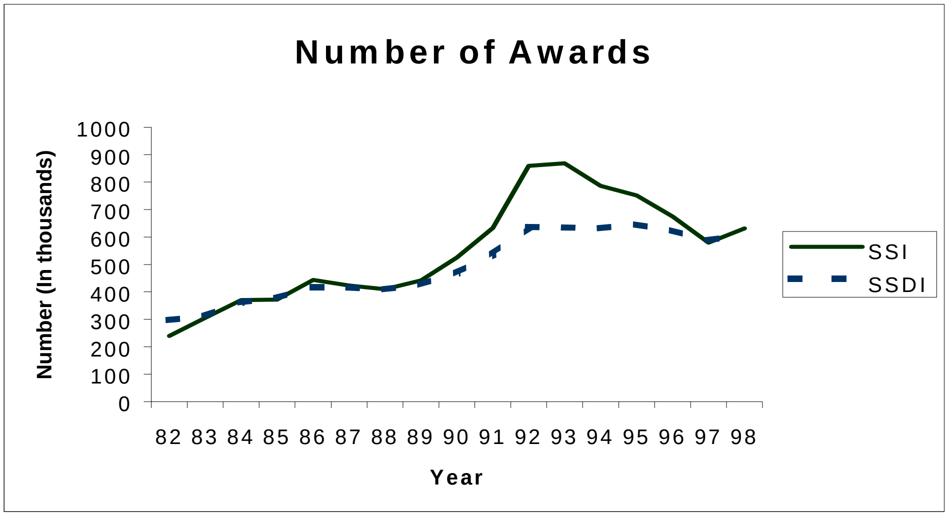
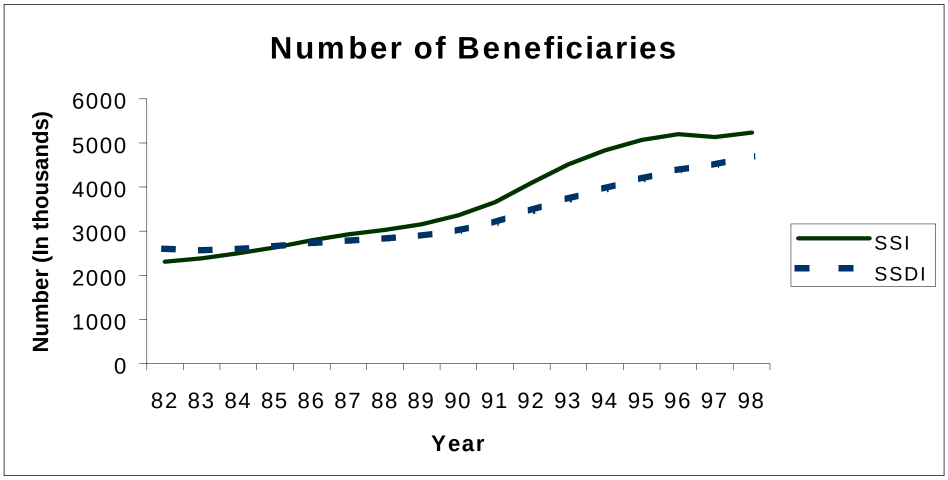


Figure 3 represents the number of beneficiaries currently enrolled in the SSI and SSDI programs. There is not a substantial difference between the number of beneficiaries or the rates of increase in this picture. However, the number of SSI recipients did grow at a slightly faster rate than the number of SSDI recipients. Since the mid-90s the reduction in SSI awards caused the annual growth of beneficiaries to more closely resemble SSDI.

Figure 3. Number of Beneficiaries



The award and beneficiary graphs call attention to the differences in SSI and SSDI programs. The programs share identical disability determination processes but are targeted to different populations. SSDI serves disabled individuals with significant work histories who have contributed substantial payroll taxes to Social Security in the past. SSI serves disabled individuals without a significant work history who meet strict income and asset limits. Children can receive SSI benefits, but they are allowed to receive SSDI benefits only as dependents.

Thus, the programs serve different clientele. Generally SSI beneficiaries are younger, more likely female, more likely minority, less well-educated, and come from single-parent families or families with lower income. Clearly, the type of beneficiary will affect the probability of entering a benefit program, the duration of time spent in the program, and the ability to return to work. So, there are a number of differences between SSDI and SSI beneficiaries, and these differences will need to be identified by data collection, used in modeling to answer the research questions, and considered in the interpretation and application of study results.

PERSONAL CHARACTERISTICS

The research literature includes several personal characteristics as possible determinants of entry and exit to benefit programs. Although most of these variables should be incorporated into models involving both SSI and SSDI, they do affect these two benefit types differently. For example, women are more likely to be awarded SSI benefits. The percentages of female recipients are 55 percent and 40 percent for SSI and SSDI respectively. Historically, women have been less likely to work than men, and therefore are less likely to qualify for SSDI. The gender inequality of labor force participation is changing, however, so the receipt of recent awards by men and women are more similar for SSI and SSDI than of past awards.

The racial distribution differs between SSI and SSDI beneficiaries also. A total of 30 percent of SSI beneficiaries are African-American and 9 percent are of other minority groups. The percentages are 18 percent and 6 percent for SSDI recipients. African Americans and other minorities are more likely to be poor and less likely fully employed as compared to Caucasians. Therefore, the racial distribution of SSI and SSDI is not surprising.

DI recipients are limited in age. Anyone reaching age 65 with a work history moves off of Social Security disability benefits to retirement benefits. Children under age 18 generally do not have a work history sufficient to qualify for SSDI benefits. Thus, potential recipients are limited to those between the ages of 18 and 64. SSI benefits are given to three groups: the aged, blind, and disabled. To qualify for aged benefits one must be over age 64 but no age limitation is given for blind and disabled recipients. The group of aged recipients is excluded from the sample in most comparisons with SSDI. Children can also receive SSI if they have a disability. About 17 percent of SSI recipients in 1997 were under age 18. Nearly 60 percent of SSDI recipients are between the ages of 50 and 64 while only 38 percent of SSI recipients are 50 or older. Overall, the average SSI beneficiary is about eight years younger than the average SSDI beneficiary.[2]

Rupp and Stapleton (1995) argue that demographic shifts in the recipient populations, specifically the increasing number of children on SSI, will lead to long term increases in the growth rates of SSI caseloads. Rupp and Scott (1995, and 1996) emphasize this same point in related articles. Relative to adults on SSI or adults on DI, children who receive SSI will have, on average, longer initial spells and

[2] 1994 NHIS data from family resources section.

longer expected life time disability stays. Thus the growth rate in SSI caseloads will likely accelerate, as children account for a larger fraction of new recipients.

Educational achievement also varies between the programs. As SSI recipients are younger and less wealthy, they are also less educated in general. Over half (52 percent) have not completed high school. Only about 3 percent of SSI recipients have graduated from college while nearly 6 percent of SSDI recipients are college graduates.

Different impairments are also prevalent among SSI recipients, when compared to SSDI recipients as seen in Table 2. While over half of SSI beneficiaries have a mental condition, either mental illness or cognitive impairment, SSDI beneficiaries have a relatively higher percentage of injuries and systemic diseases such as circulatory and musculoskeletal.[3] The differences in health condition also have a large impact on the probability of entering and exiting the benefit roles, and should lead to a difference in modeling the two programs. Oi (1998) focuses on the existence of heterogeneity among the disabled and stresses that the ability to work depends heavily on condition and severity of disability.

Table 2. SSI and SSDI Recipients by Diagnostic Condition

	1997	
Diagnostic Group:	SSI	SSDI
Infectious and parasitic diseases	1.8	2.1
Neoplasms	1.5	3.0
Endocrine, nutritional, and metabolic diseases	4.5	4.9
Diseases of blood and blood-forming organs	0.7	0.2
Other mental diseases	31.2	25.6
Mental retardation	27.7	5.2
Diseases of:		
Nervous system and sense organs	9.2	9.7
Circulatory system	4.9	12.1
Respiratory system	2.8	3.6
Digestive system	0.8	1.4
Genitourinary system	1.0	1.6
Skin and subcutaneous tissue	0.2	0.2
Musculoskeletal system	7.5	22.4
Congenital anomalies	1.5	0.4
Injuries	2.5	5.6
Other	2.2	2.0

[3] Statistics from 1998 Annual Statistical Supplement to the *Social Security Bulletin*.

Family structure also affects SSI and SSDI incidence rates as it is a proxy for household income and an indicator of support received. Rupp and Stapleton hypothesize that single-parent households positively affect the number of SSI applications. Single-parent families do not have the security of another potential wage earner if an impairment occurs. In fact, we find that SSDI recipients are much more likely to be married than are SSI recipients (42 percent and 24 percent respectively).

Large families are more frequently associated with poverty status. As the number of children increases, the probability of SSI receipt also increases. However, large families also require more income to sustain. As financial responsibilities increase, work looks more attractive than disability benefits as a job provides greater income. Statistics indicate that more SSI recipients have large families than SSDI recipients, as 11 percent of SSI recipients have three or more children under age 18 living in their household, while only 3 percent of SSDI recipients have that number of children.[4]

Self perception may also play a role in determining SSI and SSDI application rates. Not only do income, health, and economic conditions affect the decision to apply for benefits, but also one's own belief in his/her health, income, and the economy. SSDI recipients are slightly more likely to report poor health than SSI recipients are, as 62 percent of SSDI respondents in the 1994 National Health Interview Survey (NHIS) reported fair or poor health while only 55 percent of SSI respondents did so. Choi (1998) uses the New Beneficiary Survey and Social Security administrative records to explore whether subjective perceptions affect SSI application. He finds that self-reporting poor health and worrying about finances generally increases the probability of applying for SSI. Greenblum and Bye (1987) state that SSDI recipient's return to work is strongly influenced by the person's "self-assessment of capacities in relation to prior job requirements." Thus, self-evaluation may be a factor for entry and exit of both SSDI and SSI.

WORK EXPERIENCE

The work history of SSDI recipients is obviously more complete than that of SSI recipients. Since work experience is a requirement for receipt of SSDI, all beneficiaries have had extensive work experience, while 21 percent of SSI beneficiaries have never worked before (Scott, 1992). Those that have worked in the past have disrupted employment histories, as 67 percent of SSI recipients have

[4] 1998 CPS data. Sample includes 1719.67 (weighted) SSI recipients and 1605.37 DI recipients aged 18 to 64.

experienced at least one six-month or longer period without work since their first job. More surprisingly, 84 percent of rejected SSI applicants have had such a gap in their work history. Furthermore, SSI recipients have spent only 33 percent of the years since their first job engaged in jobs held for at least six months.[5]

Because of the inconsistent work experience of SSI recipients, many factors which contribute to a SSDI recipient's decision to apply for benefits may not be appropriate for SSI recipients. Labor force attachment can be measured in various ways. Muller, Scott, and Bye (1996) attempt to measure attachment by average annual pre-disability earnings in their model of probability of work while enrolled in SSI. They find inconclusive results. Earnings have a positive significant effect on the non-developmentally disabled only. Increased earnings, hence attachment, decreases the probability of work for retarded individuals while on the roles.

Burkhauser, Butler, and Kim (1995) conclude that labor force participation depends on job tenure and worker accommodation. They rely on a hazard model that measures the likelihood of a worker leaving employment in each year following the onset of a health impairment which affects work. Although these variables are relevant to analyzing SSDI beneficiaries, they do not apply to the majority of SSI applicants. Many SSI applicants have had a limiting condition their entire lives, while others were not employed at the time of onset. Thus, in studying characteristics of the job at the point of onset of disability we can gain valuable information for SSDI applicants, but not for SSI applicants.

Greenblum and Bye (1987) conclude that SSDI beneficiaries have work value levels similar to non-beneficiaries. They measure work value from three questions adapted from Goodwin. The statements, which can be agreed or disagreed with, include: "if you don't have a job, you don't feel right; a person should work on a job in order to keep the respect of family and friends; you really can't think well of yourself unless you have a job." A person with ten years out of the labor force will view these statements differently than a person who was consistently employed or searching for a job in the past decade. A high value placed on work probably contributes to the likelihood of returning to work once on the benefit roles, and it probably decreases the probability of entering the roles to begin with.

Foregone earnings and earnings risk are also likely to be more relevant for SSDI rates than for SSI. SSI recipients receive a larger percentage of their income from non-work sources. Scott (1992)

---

[5] Statistics generated from SIPP 1993, wave 1 topical module data.

finds that the majority, 72%, of SSI recipients (over age 18) earned less than $5,000 per year prior to joining the roles. Only 20% of SSDI recipients had a pre-disability income under $5,000 (Muller, 1992). Thus, foregone earnings are less important to SSI beneficiaries than SSDI beneficiaries. Kreider (1998) studies the effect that earnings risk has on participation in the SSDI program. He measures wage uncertainty by the degree of volatility in wages for each occupation, and estimates that the SSDI application rate is about 15% higher in the absence of earnings risk. The change in the application rate for SSI is likely to be much lower when controlling for earnings risk. Perhaps if we were to consider earnings uncertainty as measured by volatility in employment status than this variable may be more applicable to participation in SSI.

OTHER INCOME

It is also important to consider other sources of income, savings, and access to health insurance when examining the decision to apply for disability benefits. SSI recipients are much more likely to be in poverty, and thus, more likely to receive public assistance or welfare to supplement their income. A total of 55 percent of SSI recipients from the 1994 NHIS data fall below the poverty threshold. Only 27 percent of SSDI recipients could be considered poor by this standard. Since they are not as likely to receive additional aid, welfare variables are likely to be insignificant when included in the DI models. However, as a result of a relatively consistent work history, SSDI recipients are more likely to have accumulated savings or have significant retirement accounts, or other assets. They also have greater access than SSI recipients to employer benefits such as health insurance.

PROGRAM CHARACTERISTICS

Other benefits which accompany SSI and SSDI may have an effect on participation rates. While SSI recipients are eligible for Medicaid, SSDI recipients are covered by Medicare after 24 months. In addition, most SSI recipients are eligible for food stamps, as well as a range of state and local services. Changes in such additional programs may affect changes in entry and exit rates for SSDI or SSI.

Yelowitz (1997) considers how the rising value of Medicaid coverage may explain the recent growth in SSI caseloads in a time when access to medical coverage has been an increasing problem for persons with tenuous employer connections. By analyzing data from the Current Population Survey (CPS) from 1988-1994, he concludes that 20% of the recent growth in SSI caseloads can be accounted for by the increase in the value of Medicaid insurance. Although Medicaid is a federal program with certain minimum coverage requirements, states are free to offer a number of optional benefits. This flexibility allows Yelowitz to observe how variation in a particular state's benefit coverage over time is associated with variation in state specific SSI caseloads, while controlling for other relevant factors. Of particular interest to the policy maker is how an increase or decrease in the generosity of SSI benefits affects the labor force participation behavior of individuals who are likely to meet the disability criteria established by the Social Security determination process.

Kreider explains that a "rejection risk effect" may influence a worker's decision to apply for public transfers if program acceptance is uncertain and rejection is costly. One cost of applying for benefits is the imposed waiting period. A five-month waiting period is imposed on SSDI applicants, but not on SSI applicants. Frequently, losses in job skills and losses due to not learning new skills are incurred (Parsons, 1991). Employers may look upon rejected applicants unfavorably. They may believe that such workers might present an increased productivity risk (Aarts and deJong, 1990) and have a decreased labor force attachment (Halpern and Hausman, 1986). These costs are most harmful to SSDI applicants, as they have more regular work experience and more to lose while being out of work. SSI applicants face less change between pre-application and the waiting period.

Knowledge of disability programs also affects the decision to apply for benefits. Application rates would be low if SSI and SSDI are unfamiliar programs to the disabled population. Choi (1998), Coe (1985), and Hill (1990) each find that over half of the disabled assert that lack of information or misinformation is a reason for not applying to SSI. Choi bases his findings on data from the New Beneficiary Survey while both Coe and Hill use the Panel Study of Income Dynamics for their analysis. Unfortunately, lack of information regarding SSDI benefits has not been previously considered as it has been for SSI. Thus, we do not have the data to determine whether the disabled population is more familiar with SSDI or SSI benefits. The DES should provide answers to this question.

Work incentives also differ somewhat between SSI and SSDI programs. SSI recipients are allowed to have earnings above the substantial gainful activity (SGA) level as long as their income does not exceed the federal benefit rate. Plans for achieving self support (PASS) also allows SSI recipients to deduct work expenses related to their impairment from their income when determining eligibility. The SSDI program offers a trial work period in which a recipient may engage in SGA and still earn benefits for up to nine months in a 60-consecutive time period.

Although SSI recipients are less likely to have worked prior to receiving benefits, they are more likely to have work experience after joining the benefit roles. A total of 22 percent of working-age SSI recipients report having received some earnings after the year in which they applied while only 12 percent of SSDI beneficiaries report having post-application earnings (Scott, Muller). These figures are self-reported and SSDI beneficiaries have a greater incentive (due to a greater opportunity cost) to not claim their post-application earnings. However, SSI recipients are also more likely to be referred for vocational rehabilitation (VR) services as SSDI recipients (186,826 and 45,248 VR referrals, respectively, in 1993). Nevertheless, outcomes of VR are not as favorable for SSI recipients. Only 1 percent of referred SSI recipients are successfully rehabilitated from VR agencies while nearly 5 percent of referred SSDI recipients are rehabilitated (Mashaw, 1990).

MACROECONOMIC ISSUES

Rupp and Stapleton list several economic variables that may affect participation in disability benefit programs. Benefit awards seem to increase along with declines in the labor force participation rate. Such a variable may have an effect on both SSI and SSDI, however, the effect on SSDI would be enhanced due to the long-term effect of women's increased coverage by disability insurance. As the poverty rate increases the number of SSI applicants also increase. However, the number if SSDI applicants should not be affected by such a change. Indicators of the business cycle such as the unemployment rate also may be included in the models. Several authors include binary variables representing each year to accommodate such effects. Rupp and Stapleton find in their study that the 1990-1991 recession did play a significant role in the growth rate of both SSI and SSDI.

CONCLUSIONS

It is imperative to recognize the differences between SSDI and SSI claimants when creating models of benefit participation. SSI recipients do not have a consistent work history as SSDI recipients do. Thus, variables relating to work such as earnings, accommodation, work values, employee retirement accounts, and unemployment rates will not be as important in explaining SSI entrance. SSI caseloads have been affected by a number of other factors – the growing value of Medicaid, pressures exerted by states to shift disabled welfare recipients to SSI, and a large increase in the number of children entering the program – that do not appear to bear on the SSDI program. Yet, most published work on disability by economists has treated these two populations as interchangeable. While it is true that they all meet the same definition of "disability," they face very different economic and non-economic alternatives.

Thus, it is mandatory that SSDI and SSI claimants be examined separately for changes in behavior as a result of disability. The major reason why previous researchers have not made this distinction is that the data sources are not sufficiently precise, nor do they contain enough persons with disabilities, to permit an analysis of the differences between these two populations. The DES will provide a large national sample of persons with disabilities. Further, it will gather the most detailed data ever on the physical, mental, and functional status of those individuals. The Westat analysis and evaluation team will utilize these two features of the DES to study and model the similarities and differences between SSDI and SSI populations.

II. SSA Core Research Objectives

This section presents our detailed analysis plans for the core research objectives specified by SSA. The structure of the section follows the research objectives, as presented in the introduction. The four objectives are:

1. Providing an estimate of the U.S. population of persons potentially eligible to receive disability benefits;

2. Determining how many persons in the U.S. population do not meet eligibility requirements today, but may meet them at some point in the future;

3. Identifying factors that enable disabled individuals to remain in the work force (e.g., treatments, accommodations, and interventions);

4. Identifying self-reported measures appropriate to estimating future program eligibility

A. Objective 1 - Providing an Estimate of the U.S. Population of Persons Potentially Eligible to Receive Disability Benefits

INTRODUCTION

SSA recognizes that there are a substantial number of people in the U.S. population who have severe impairments but who are not currently receiving disability benefits. Some are working now, but, for various reasons (e.g., a downturn in the economy or a further progression of their impairment), might soon decide to stop working and apply for benefits. Some are currently not working, but, nevertheless, have not applied for benefits. The DES has been designed to provide an understanding of the size and characteristics of the group that may potentially receive disability benefits. This section of the proposal describes how we will identify this group for future policy analysis.

To determine how many individuals in the population are “disabled according to definition,” the DES will analyze the physical, mental, and functional status of a sample of approximately 5,500 individuals. The sample strata will be representative of those in the general population who are severely disabled, moderately disabled, or not disabled, plus a sample of current SSDI and SSI beneficiaries. But how does one decide who is disabled and who is not, if they have not actually applied for SSDI or SSI benefits?

The DES will use the SSA definition of disability and the Disability Determination Services (DDS) will “simulate” the current decisions in the various states, but without regard to assets or current work status. In addition, Westat will develop a “gold standard” of disability determination as a standard of comparison. This will be a state-of-the-art process constrained only by the physical requirements of the Mobile Examination Centers, the cooperation and safety of the respondents, and the limits of our medical and diagnostic knowledge. The major thrust of the DES is to acquire this disability defining information about the targeted populations. The information will be used to estimate how many people are potentially eligible to receive disability benefits.

DEVELOPMENT OF A GOLD STANDARD FOR DISABILITY

We begin this section with a discussion of the “gold standard” for disability determination; what it is, and the process Westat will use to develop it. For SSA purposes, a person is disabled if, using the definition of disability in the Social Security Act, the regulations for determining disability, and the current disability determination process, an individual is found eligible for disability benefits. There is no authoritative external source which identifies a person as truly disabled. Therefore, in order to answer some of the research questions in the DES, it will be necessary to develop what SSA calls an “achievable” gold standard.

The achievable gold standard is expected to provide the SSA and Westat with a standard for disability against which to compare disability decisions made in a variety of other ways. In the fourth core objective, self-reported measures will be evaluated as a replacement for a full-scale medical study such as the DES to estimate program eligibility. The results of using self-reported measures will be compared to the gold standard as well as to the outcome of the simulated disability determination process. There are two redesign objectives which also will utilize the gold standard. To test the validity of the redesign process, Westat will compare the results of a simulated redesign process with the gold standard as well as the current DDS process. In addition, one of the redesign objectives is to identify self-reported measures appropriate to estimating program eligibility. Again, the disability results obtained with self-reported measures will be compared against the gold standard.

There are three operational decisions that need to be made in developing a gold standard: (1) the information to use in making a decision; (2) the individuals or groups of individuals to use to make decisions; and (3) the process to use to make decisions. Regarding the first decision, we recommend that the data to be used are drawn from information collected in the DES. Table 3 lists the type of information collected in the DES at each stage of the study. While it might be presumed that a lot of information is likelier to provide a better gold standard than a small amount of information, it would, nevertheless, be cumbersome and impractical to use all the information collected in the DES. In addition, some types of information will be ascertained in varying levels of depth, by different people, and in different formats (e.g., self-report and observation). Thus, a subset of DES data seems appropriate.

Therefore, as part of developing a gold standard, it is proposed that the data collected in the DES be pared down to a more manageable dataset using an agreed set of criteria for eliminating some information. For example, decisions would need to be made on whether observational data were preferable to self-reports, or data from the Initial Screener (possibly from a proxy reporter) were preferable to data from the Follow-up Screener (from the individual respondent). The goal of the process is a reduced dataset that will be used in the development of the gold standard. As part of the medical examination protocol, physicians, nurse practitioners, and psychiatric social workers will provide their assessment of the subject person's (SP's) residual functional capacity. Since their opinion will be based on face-to-face contact with the SP, it is important that it be included in whatever pared down dataset is constructed. This dataset will then be used for developing the gold standard.

Those who develop the gold standard will have expertise in the medical, functional, and evaluation aspects of disability (e.g., physical medicine, psychiatry and/or mental health, vocational rehabilitation, functional assessment, job placement and work requirements). Westat will develop a "gold standard panel" comprised of such experts. Using the opinion of the interviewer, physician, nurse practitioner, and psychiatric social worker, as well as all other information in the reduced dataset, the panel will make disability determinations on each SP based on SSA/Westat guidelines for labeling a person as "disabled" or "non-disabled."

However, the key question of how many persons are potentially eligible to receive SSA disability benefits today will be answered by a direct simulation of the current SSA/DDS disability determination decision process. The DES respondents will be evaluated for disability benefits by the DDS claims officer, ignoring their current work status, i.e., as if they have applied for benefits. Thus, core objective 1 involves applying the simulated current disability determination process to the DES sample. When the sample is inflated to represent the population from which it is drawn, the result is an estimate of the population potentially eligible for SSDI and SSI benefits, based solely on physical, mental, and functional status, and ignoring their work or asset situation.

This section discusses using simulation methods to emulate the disability determination decisions that must be made when an individual applies for disability benefits. Estimation issues are discussed later (section II-B) in the context of estimating future eligible populations.

Table 3. Information Collected for Each Stage of DES

Telephone Interview n = 95,238	In-person Interview n = 11,111	Medical Examination n = 5,500	Medical Evidence of Record n = 5,500
Initial Screener	Follow-up Screener	Clinical Examination	Follow-up with Testing Physicians
1. Household Enumeration	1. Household Enumeration	1. Performance Assessments	1. Test reports (e.g., x- rays, pathology) from treating physicians and health care facilities
2. Functional Limitations - communication - sensory - lower extremity functioning -upper extremity functioning - interpersonal	2. Functional Limitation Scales aid in identifying - mental illness - cognitive deficit - physical impairment	2. Tests and Results - mobile examination center tests (MEC) - outside referrals	
3. Health Status - selected impairments and assistive devices - perceived health status	3. Diagnostic Tests Composite International Diagnostic Interview (CIDI) - Mini-mental State Exam (MMSE)		
4. ADL/IADLs	4. ADL/IADLs		
5. Current SSDI/SSI Status	5. Performance Measures 19 standardized tests to confirm function limitations		
6. Work Status	6. Perceived Disability		
	7. CSI - health behaviors - medical history - access to health, educ., and vocational services - social and community living - work status, history, and environment - program knowledge - economic resources		

<u>Initial Screener</u>	<u>Follow-up Screener</u>	<u>Clinical Examination</u>	Follow-up with <u>Testing Physicians</u>
purpose:	purpose:	purpose:	purpose:
- Increase sampling efficiency - Analyze proposed cut-points between severe and moderate disability - Allocate individuals to four groups - Describe the population screened and compare with other national surveys	- Confirmation of severe cognitive and mental impairment and the functional conse-quences of physical impairments - Delineate persons with borderline impair- ments from those with severe impairments - Confirm SSI/SSDI beneficiary status	- Creates direct clinical observations and judgments about study participants - Review reported signs, symptoms, and medical conditions - Provide information for the disability determination process	- Confirms clinical observations of MEC staff - Protects SP from danger and discomfort of retest - Provides information for the disability determination process

THE DISABILITY DETERMINATION PROCESS

The number of disabled beneficiaries is determined by five distinct factors: 1) the onset of impairment, 2) the decision to apply for benefits, 3) the determination of eligibility, 4) the appeal of denials, and 5) the decision to return to work. Accurate determination of the number of potential disabled beneficiaries requires an understanding of the trends in the general health of the population (factor 1) as well as the various decision points in the determination and appeal process (factors 2-4). In addition, increasing attention is being given to the potential decision on the part of beneficiaries to forego their benefits and return to work (factor 5). A convenient way to conceptualize the various components that contribute to the number of beneficiaries and to understand their relative contributions is to construct a definitional equation that relates beneficiaries to population. These components, and rough probability estimates taken from various studies, are combined in the following equation:

disabled beneficiaries = population x

{[disabled/population]*[applied/disabled]}*{[accepted/applied]+
0.07 0.75-0.85 0.50
[accepted/appealed]*[denied/applied]*[appealed/denied]- [return to work/accepted]}.
0.68 0.50 0.65 .005

That is, the number of disabled beneficiaries is a product of the population times the disabled to population ratio times the probability of application times the percentage who apply that are accepted for benefits, plus the percentage who successfully appeal, minus the percentage who return to work.

In order to perform useful simulations with such models, it is necessary to adequately specify the process by which a working individual moves from the status of not disabled to disabled, to application for disability benefits, to allowance of benefits, or if denied, to the appeal process. To accurately estimate the probability of transition from one status to the next, it is vital that the appropriate sample of the population is available for analysis and that the relevant information is collected about people in these samples.

The SSA currently determines the eligibility of applicants for SSDI benefits through a five-step screening process. Each step is based on a different set of information.

(1) *Earnings screen*. Individuals are denied benefits if they engage in substantial gainful activity (SGA). Activities are substantial if they entail significant mental or physical activity. They are considered gainful if they are performed for pay or profit. Applicants earning more than $500 per month are not passed to the next screen. For the purposes of the simulated disability determination in the DES, we will ignore this screen.

(2) *Severity screen*. An applicant's impairment is considered severe if the disability will either lead to death or last at least 12 months. An impairment must be determined severe before the applicant is passed on to the next step.

(3) *Medical screen (listings of impairments).* Applicants are evaluated as to whether they "meet" or "equal" the list of impairments. Currently, the codified criteria include more than 100 impairments. Applicants with impairments that meet these criteria are allowed benefits without further evaluation of their claim (after a waiting period of 5 months has been met). To meet these criteria, an applicant's impairment must present the symptoms, signs, and laboratory findings defined for one of the impairments on the list. If an applicant has an impairment that is not on the list, but which is considered medically equivalent and equally severe to a listed impairment, then the claim is also allowed.

4) *Capacity for Past Work.* For those applicants not meeting step 3, they are judged on their residual functional capacity to determine whether they meet the requirements of the main jobs they have held in the past. If they are deemed able to perform past work, they are denied benefits. The remaining applicants are passed on to step 5. The evaluation of work-related limitations in this step is different from that in step 2. In step 2, applicants were judged on the presence of such limitations, such as the ability to lift or to understand basic instructions. In this step, the extent of such limitations is assessed to compare the applicant's residual capacity with the demands of jobs they have previously held.

5) *Capacity for Other Work.* Those applicants who are judged incapable of performing their past jobs must also be scrutinized according to their ability to perform any other job in the national economy. To make this judgement, a vocational grid is used to determine whether the applicant can work in employment consistent with his or her residual capacity.

Lahiri, et al. (1995) modeled disability determination as a sequential model incorporating each of the steps listed above. The purpose of their multistage model is to capture the structure of the process so that proposed policy changes can be simulated, as well as estimate the pool of eligibles in the general population. Such a model is useful in forecasting the future pool of eligibles and thus the trend in SSDI and SSI caseloads.

Each step of the determination process requires a different set of data. For instance, step two (severity of claimed medical impairments) deals with functional limitations, whereas step three (impairment list) is based primarily on medical evidence. Lahiri and others include all the variables used in the estimation of each stage of the determination process in a reduced form estimate. They found, as expected, that the reduced form model had a good deal of predictive power. But only a relatively small number of the explanatory variables had reasonable statistical confidence levels.

Therefore, a reduced-form modeling approach may be appropriate for predictions of future eligibles and caseloads, but it appears that it is not sufficient to understand the effects of specific factors in the determination process. For the purposes of the DES, it is appropriate to pursue both approaches. We plan to pursue various specifications of structural and reduced-form models in order to predict future applications and awards. However, to estimate how many people are eligible today for disability benefits we will use direct simulation methods.

SIMULATING THE DDS PROCESS USING REVIEWERS OF CLAIM FOLDERS

We plan to use actual DDS examiners to determine the eligibility of hypothetical applicants. The DES "simulated" claim folders will approximate those of actual cases that are submitted to DDS decision-makers by SSA. Using reviewers to simulate the appeal process raises several concerns. These difficulties include: the ability to assemble such a panel with a representative experience base; the review of case folders that will in several ways be more complete, and in other ways be less compatible, with actual case folders normally reviewed at the ALJ level; and the absence of face-to-face interviews of the SPs.

In addition to implementing a simulation of disability determination decisions, it will also be necessary to evaluate the simulation process, both in the pilot and the main study. To evaluate the simulation in the main study, two questions will be answered:

1. Are the DES simulated claim folders typical of the claim folders compiled by the DDS?

2. How reliable are the disability decisions made by decisionmakers selected to review the simulated claim folders?

Our approach to answering these questions is discussed below.

1. Are the DES simulated claim folders typical of the claim folders compiled by the DDS?

Despite our best efforts, there will be some differences between the process used to create the simulated disability claim folders and claim folders created under the actual system implemented through the DDS offices and FDDS.

To ensure that these differences do not become unacceptably large, Westat proposes to ask SSA to retrieve the original records for a sample of those SPs who have applied for benefits. These original records will be qualitatively compared with the DES simulated claim folders to judge the overall adequacy of the folders used in the simulated disability determination. These folders will be evaluated in a side-by-side review to judge the completeness of the folders, the number of test results included in the folders, and items included in common and not in common.

Additional criteria for this review will be developed prior to the pilot study. For example, Westat proposes retrieval of only those folders that have been more recently reviewed by SSA. This is because current beneficiaries and denied applicants may have applied many years in the past, and their health status may have changed since the original date of application. Moreover, older claim folders may be more difficult to retrieve. Criteria for the review will be developed during the pretesting, will be evaluated in the pilot test, and implemented in the main study.

2. How reliable are the disability decisions made by decisionmakers selected to review the simulated claim folders?

Reliability is the extent to which a measuring device produces the same results on multiple application to the same phenomenon (Litwin, 1995). Interrater reliability provides a measure of how well two or more evaluators agree in their assessment of a phenomenon. To assess interrater reliability in the simulation, our study design will consist of repeat determinations of a subset of simulated claim folders. Decisions will need to be made on the specific study design of the replicate study, sample size requirements, and analysis. These issues are discussed below.

(1) Study Design Options

Besides the practical criterion of ease of administration, a design is desirable if it uses the data efficiently so that statistical tests have high power or provide narrow confidence intervals. Therefore, Westat will investigate alternative study designs to assess interrater reliability. The simplest design is to assign equal numbers of folders to pairs of reviewers so that distinct reviewer pairs (claims examiners and physicians) do not examine the same folders and all folders are examined by exactly two reviewers. As discussed in Dunn (1992), other designs are available including those in which some folders are examined by more than two reviewers, and some reviewers are paired with more than one other reviewer. Many such designs are possible, although few can be practically administered. Currently it seems most practical to have two independent reviews for a sample of the 5,500 simulated claim folders. The second reviewer of a claim will complete his or her review without knowledge of whether the claim has been previously reviewed, or knowledge of the results of the previous review.

(2) Sample Size

For reliability studies with a continuous outcome, Donner and Eliasziw (1987) show that small sample sizes (fewer than 20 reviewers and fewer than 50 subjects per reviewer) give high power to test the hypothesis of non-zero intraclass correlation, a measure of reliability for such data. While accurate estimation of Cohen's Kappa (the discrete-data analogue of intraclass correlation) is more difficult than testing the hypothesis of zero Kappa, it may still suffice to use a far smaller sample size than the maximum number available.

To understand the implications of using fewer than 5,500 claims in the interrater reliability study, we calculated the precision with which Kappa could be estimated in an experiment with a pair of reviewers per folder, but with alternative choices of the number of folders to be examined by each pair. If a reviewer's assessments of different folders are independent, then the number of reviewer pairs is irrelevant to the calculation; only the number of folders is needed. Thus, we can imagine that there is only a single pair of reviewers who generates one 2 X 2 cross-tabulation of their decision to allow or deny.

We present here only a rough calculation of the precision that would be attached to an estimate of Kappa, and how that would impact on the number of folders to be reviewed. The calculation is based on computing a confidence interval for Kappa and observing the impact that varying sample size has on the width of that interval. A good choice of sample size is one that makes the interval neither too wide (low precision) nor too narrow (unnecessarily high precision).

Two hypothetical cases of concordance and discordance will be considered, as shown in Table 4. Kappa values of 0.6 and 0.8 are used indicating moderate, and good levels of concordance, respectively. In each table, the marginal proportions (the stipulated eligibility status of the claim folders being rated) are 50 percent for allowance and 50 percent for denial.

The width of the confidence interval depends upon the true value of Kappa and the sample size. A width of ±0.10 is attained when there are about 1,500, 1,000 and 600 folders for poor, moderate and good Kappa, respectively (see Table 5). Thus, it would appear that a sample of less than 1,000 folders will need to be reviewed.

(3) Analysis of Interrater Reliability

The purpose of the interrater reliability experiment is to estimate a statistically appropriate measure of agreement that will summarize the extent to which ideal agreement is achieved. Simple percent agreement is not an appropriate measure because, merely by chance, this proportion can be extremely high. For example, suppose that 50 percent of claims are allowed and 50 percent are denied. If two reviewers independently made random guesses with these proportions, there would still be 68% agreement simply by chance.

Table 4. Two Hypothetical Cases of Concordance and Discordance of Reviewers

Moderate Concordance (Kappa = 0.6)

	Reviewer 2	
Reviewer 1	Allow (0.5)	Deny (0.5)
Allow (0.5)	0.40	0.10
Deny (0.5)	0.10	0.40

Good Concordance (Kappa = 0.8)

	Allow (0.5)	Deny (0.5)
Allow (0.5)	0.45	0.05
Deny (0.5)	0.05	0.45

Table 5. 95 Percent Confidence Interval Width for Specified Total Number of Folders

	95% Confidence Interval Width	
	Kappa = 0.6 (moderate)	Kappa = 0.8 (good)
Number of Folders		
100	±0.314	±0.235
200	0.222	0.166
400	0.157	0.118
600	0.128	0.096
1000	0.099	0.074

Overview of relevant statistical methods. The primary methods of analysis are the log-linear model to test concordance, tests of the hypothesis of marginal homogeneity, and Cohen's Kappa statistic as a measure of agreement. Suppose that there are two reviewers for each replicated folder. The results of their determinations can be displayed in a 2 X 2 table (Table 6).

Table 6. Results of Simulation in a 2x2 Table

	Reviewer 1		
Reviewer 2	Allowed	Denied	
Allowed	a	B	P_{2+}
Denied	c	D	P_{2-}
	P_{1+}	P_{1-}	

Agreement between the reviewers is considered good when the observed total number of folders in the table's diagonal cells (denoted a and b) is significantly higher than the expected number under the hypothesis of independence between the two reviewers (which would imply complete absence of agreement). A log-linear model provides a way to compute the expected number of agreements under independence and a way to assign a p-value to the difference between the observed and expected numbers (Landis and Koch, 1977; Velema et al., 1991). The models are general and can be extended to more complex designs, such as those with heterogeneous strata or with many reviewers who are not always paired and not all of whom examine the same folders. Heterogeneity might be due to the state in which the disability claim originated, to other claimant features such as age, gender, education and race, or to features of the claim itself (e.g., body system affected or the claimant's functional ability to perform certain tasks). For such situations, the null hypothesis of independence between reviewers is more complex than in the table above, but the log-linear model can be used to estimate expected counts under independence.

In the above table, Reviewer 1 assigns the two possible labels to claims in the proportions p_{1+} and p_{1-}—denoting allowance and denial, respectively, and summing to 100 percent. These categories may be viewed as a reasonable ranking of the likelihood of a successful claim. Similarly, Reviewer 2 assigns the labels in the proportions p_{2+} and p_{2-}. Although not sufficient, it is necessary that the two sets of proportions be equal if there is to be interrater agreement. With such paired data, this is called the hypothesis of marginal homogeneity.

When the simplest design is used and data are in the form of the 2 X 2 table, the McNemar test of marginal homogeneity is an appropriate statistic to use for analysis. If, however, the design is more complex than a single two-way table (i.e., if there are additional sources of variation), the generality of the log-linear model is again needed. For this, we can turn to the solution of Lipsitz, et al. (1990) who showed how marginal homogeneity could be tested by a log-linear model.

The log-linear model provides a test of agreement. In contrast to this, Cohen's Kappa statistic is a chance-adjusted measure of agreement. In general, the Kappa statistic takes the form

κ = (observed agreement - expected agreement)/(1 - expected agreement)

so that it is the ratio of the observed improvement over expectation (under the hypothesis that reviewers are independent) to the largest possible improvement over expectation. The analyst has the choice of whether or not to take into account the normal ordering of the outcome (see Brennan, et al., 1981 and Graham and Jackson, 1993 for extensive discussions of κ). Since κ measures improvement over expectation under the null hypothesis of no agreement, and the log-linear model provides the estimates under this null hypothesis, the two approaches are linked; if a log-linear model can be applied to the data, then κ can be calculated.

MODELING THE APPEALS PROCESS

Persons with impairments who are denied benefits during the disability determination process have the opportunity to appeal their case at several levels. The first level of appeal, known as reconsideration, is conducted by the same DDS that made the initial determination but with a different team of evaluators. The second level of appeal is to an Administrative Law Judge (ALJ) of SSA. The third level is to file a request for review by the Social Security Appeals Board in Washington, D.C..

Those denied at this stage can take their case to Federal Court. For purposes of the DES, the SSA has indicated that only the ALJ level of appeal will be addressed.

According to a study by Benítez-Silva, et al. (1998), roughly half of those individuals who initially apply for benefits are denied, and 65 percent of those who are denied subsequently appeal that decision. Eventually, 46 percent of those who appeal are allowed benefits, bringing the total acceptance rate to 72 percent of those who applied. Therefore, modeling the appeal process is critical to estimating the overall determination and predicting future caseloads. Further, Benítez-Silva and others have found that estimating the appeal process separately from the determination process is important for understanding the extent and source of Type I and Type II errors in the process.

To represent the appeals process, two types of decisions must be made: which applicants will make an appeal, and what the ultimate decision will be. Given the importance of ALJ hearings in the disability determination process, it is important to properly represent ALJ decisions. However, ALJ decisions include face-to-face contact with an applicant, which will not be possible in the DES simulation. Moreover, about 75 percent of applicants at the ALJ level have legal representation. Therefore, simulating the ALJ decision would be much more challenging than the DDS decision. We do not believe it is feasible to simulate these decision, so we will use a modeling approach.

Our modeling approach will consist of determining the best predictive model for each decision (decision to appeal, ALJ decision) and then applying the model to "applicants" in the DES simulation. Benítez-Silva, et al. (1998) extended Lahiri's methodology to include modeling of the appeals process with data from the Health and Retirement Survey, but did not link their data to SSA data. We will explore the use of this method, using external datasets (e.g., the Health and Retirement Survey and NHIS-D) as well as SSA data files before DES data become available.

Examples of variables we will use are type and degree of disability, demographics, job and income information. The general characteristics are enumerated in Table 7, which lists variables employed in past studies. During the pilot phase of the DES, the available existing databases and the information that can be made available to Westat for the purposes of model development will be determined. We will seek to collaborate with current modeling efforts at SSA (Vaughan and Wixon) as well as reviewing other studies that model quasi-judicial determinations.

SSA will be asked to provide claim folders for SPs who have applied for disability benefits, regardless of the original decision or the decision made on appeal. Using the same types of methodology proposed for evaluating the interrater reliability at the first decision stage Westat proposes to compare the results using the two-stage modeling process with the true SSA determination documented in the SPs' actual case folders.

It is likely that between-state variation in allowance rates may be affected by procedures used to conduct the simulation (e.g. central review by disability examiners). If so, reliability will be spuriously increased. This can be circumvented by using an algorithm with state-specific allowance rates (based on available data and state variation estimated in the pilot study) to infer what reliability would be if the reviewers reflected interstate variation. Using the statistical bootstrap method (Efron and Tibshirani, 1993), the algorithm will give Cohen's Kappa and its confidence interval for a hypothetical data set with differential allowance rates.

Besides state variation, other sources of variation could also be handled in this manner. With its origin in simulation, the bootstrap is a computer-intensive method of estimating variability that could not be estimated by purely analytic methods. It involves sampling repeatedly, with replacement, from an actual data set, each time generating an artificial data set to which a statistical analysis is applied. The many (typically several hundred) results of these repeated "bootstrap analyses" are saved, and it is their variability that yields the key outcome such as a confidence interval for Cohen's Kappa.

Roughly speaking, the algorithm uses the statistical bootstrap to create an artificial data set that combines actual disability folder ratings and additional variation chosen from the known state-specific set of allowance rates. New terms that reflect inter-state variation are added to the log-linear model that describes the data without state-terms. Each such model yields a Cohen's Kappa and, as noted above, a few hundred such bootstrap estimates are used to form a confidence interval for Kappa.

Reliability may vary among the four strata. Indeed, it seems likely that, as severity of disability decreases (from "severely disabled nonbeneficiaries" to "minimally or nondisabled nonbeneficiaries"), reliability of assessment will also decrease. Dealing with this requires, first, that such heterogeneity is recognized and, second, that stratum-specific Kappa estimates have adequate precision. For the first issue, Barlow, Lai and Azen (1991) propose a version of Cochran's Q-statistic to test homogeneity of Kappa values. Using a chi-square with one less d.f. than the number of strata, the test indicates whether there is a common Kappa. If there is, results from strata can be pooled. If not, Kappas from distinct strata should be reported separately. Of course, the ability to recognize that Kappas differ among strata (power of the test of heterogeneity) will depend upon the number of folders within strata, with equal allocation generally being optimal.

DATA REQUIREMENTS

The transition from an individual's status before a disability occurs until the final decision on benefits is made takes place in a sequence of events. The decisions at each juncture depend upon information about an individual's current status and past activities. Therefore, data required to model and simulate the process must reflect the sequence of events; they must also include individuals who end up on each side of the decision at each critical decision point. Most importantly, they must include sufficient explanatory information about these individuals. Previous studies have used retrospective information about subjects [Health and Retirement Survey (HRS) and the New Beneficiaries Follow-up (NBF) Survey]. Some studies have matched longitudinal surveys such as the Survey of Income and Program Participation (SIPP) Survey to Social Security Administration administrative records (Form 831) in order to build a history of pertinent activities. We intend to use both approaches in the DES.

In addition to the major data collection effort to obtain retrospective information in the DES, we plan to request permission for the SSA administrative records of an individual's work history to be attached to the interview sample. One problem encountered in previous studies is poor recall of individuals when asked about earnings streams and work histories, particularly for those who have been on disability for some time. Obtaining work and earnings histories of these individuals would markedly improve the quality of the data used in the analysis, particularly that which pertains to the future caseload predictions (objective 2). Table 7 indicates the data that have been used in previous studies of SSDI. We expect to have all these items available from the DES and SSA administrative records, at least for those SPs who give their informed consent.

Table 7. Data Previously Used for Statistical Estimation

Variables	Apply	Determination Process				Appeal	RTW
		Stp 2	Stp 3	Stp 4	Stp 5		
Column	1	2	3	4	5	6	7
Expected Earnings	X						
Earnings in Preceding Year						X	
Wealth						X	
Primary Insurance Amount	X						X
State-time allowance rate	X						
State-time unemployment rate	X						
Employer Accommodation	X						X
Job Characteristics	X					X	X
Industry	X						
Union Status	X						
Employer Size	X						
Education	X					X	X
Tenure on job	X						
Hours worked in application year						X	
Age, Race, Gender	X	X				X	X
Marital Status	X	X				X	X
Medical Events and Conditions	X	X	X			X	X
Activity Limitations (ADLs, IADLs etc)		X	X			X	
Work Experience and Demands				X		X	
Time					X	X	
Vocational Grid					X		
Vocational Rehabilitation							X
Work-Incentive Provisions							X
Reasons for Working							X

Column 1: Burkhauser, Butler, Kim and Weathers (1997), “The Importance of Accommodation on the Timing of Disability Insurance Application: A Choice-Based Semiparametric Hazard Model,” Aging Studies Program Paper No. 12, Center for Policy Research, The Maxwell School, Syracuse, NY: Syracuse University, October.

Columns 2-5: Kajal Lahiri, Denton R. Vaughan, and Bernard Wixon, “Modeling SSA’s Sequential Disability Determination Process Using Matched SIPP Data,” Social Security Bulletin, 58(4), Winter 1995.

Step 2: Impairment Severity
Step 3: Meeting and Equaling the listings
Step 4: Residual functional capacity for past work
Step 5: Residual functional capacity for other work

Column 6: Hugo Benítez-Silva, Moshe Buchinsky, Hiu Man Chan, John Rust, and Sofia Sheidvasser, “An Empirical Analysis of the Social Security Disability Application, Appeal, and Award Process, Yale University, February 1998.

Column 7: Hennessey, John C. 1997. “Factors Affecting the Work Efforts of Disabled-Worker Beneficiaries.” *Social Security Bulletin,* Vol. 60, No.3, 3-20.

B. Objective 2 - Determining How Many Persons in the U.S. Population Do Not Meet the Eligibility Requirements Today, but May Meet Them Sometime in the Future

As a further means of establishing the potential size of the disability rolls, SSA would like to estimate the number of individuals who do not meet the eligibility requirements today, but might meet them sometime in the near future. This could come about because persons with moderate impairments may progress to more severe impairments, for reasons related to why people decide to apply for benefits in the first place, or because changes are made in the eligibility requirements or the disability determination process. Other factors, such as changes in the general employment situation or eligibility requirements of other income support programs might also have a significant impact on the size of the disability rolls in the future.

We believe that the most accurate and effective way of predicting future trends is to conduct a follow-up study of DES participants. This would make it possible to assess changes in severity of impairment, functioning, individual behaviors, and administrative processes between the two observations. Although a follow-up study may be preferable, meeting the research objectives is also possible using highly sophisticated modeling approaches. This assumes that matched SSA earnings data will be available to the Westat research team for most of the DES respondents. The following discussion describes our approach to this research objective.

Future trends in the SSDI and SSI programs are driven by several forces:

- Exogenous events related to accident and morbidity trends that determine the number of persons with specific impairments;
- Social and individual expectations (partly shaped by policy) that help to determine personal and collective attitudes toward impairment and disability;
- Economic conditions and changes in policies determining eligibility and level of benefits from other income-support programs; and
- The propensity to apply for SSDI and SSI and the eligibility process criteria and level of support of the SSDI and SSI programs.

The first three sets of factors relate to events outside the policy parameters of the SSDI and SSI programs and are determined primarily by environmental, demographic, and economic factors that typically change only slowly over time[6]. For example, accident rates are tied to the nation's industrial composition and the conscious efforts by management and labor, supported by the intervention of state and federal programs, to improve workplace safety. These trends are also shaped by the demographics of the workforce, in particular the age distribution of workers within various industries.

Attitudes toward disability (e.g., efforts to accommodate disabilities in the workplace) and policies of other income-support programs (e.g., General Assistance and AFDC to TANF) also change over time. To account for changes in these factors, we will engage in secondary analysis of established data sets that are collected on a regular basis to trace the course of key variables related to the first three sets of factors.

Data assembled in the DES will provide a cross-sectional representation of the U.S. population. We propose to use the information gathered in the DES to model the decisions of persons with disabilities, as well as the administrative determinations of the disability system. These models will then be used to simulate system outcomes under alternative states of the world or different behavioral assumptions.

Thus, we intend to follow a 3-step approach to meet this objective:

- **Step 1:** Model the decision of individuals with impairments to apply for disability benefits;
- **Step 2:** Model the SSA administrative decision to allow or deny benefits (including appeals);
- **Step 3:** Apply the models to predict future program size and cost.

We propose to use two types of modeling techniques: the traditional reduced form (data reduction) technique which combines all appropriate variables in the model at one time, as well as a more sophisticated structural modeling approach. The advantage of a structural modeling approach is that it has the ability to show causal relationships among variables. The reduced form technique is only able to show correlational relationships.

We also propose to develop a structural model approach that is dynamic. Dynamic modeling is an iterative approach to structural modeling in which the results of one modeling exercise informs the

[6] Some factors may affect SSA program policy parameters if changes are made in the program in response to these trends (e.g., tightening eligibility requirements in order to slow growth rate in enrollments and program costs).

contents and process of the next modeling exercise. We propose to use it primarily to model the decision to apply for disability benefits. While we recognize that individuals do not choose to suffer impairments, we believe that individuals with impairments may have some choice about whether, or when, they apply for benefits, and to which program. Our structural models are developed and described in this section of the proposal. Once the structural models have been estimated, we will use them to make predictions about SSA program size in the future through a process of microsimulation. Microsimulation uses the properties of the models (who applies and whether their claim is allowed) to predict how the population will respond to certain stimuli. Microsimulation enables us to change those stimuli and see what the impacts are.

The remainder of this section is divided into several sub-sections. We begin with the modeling approaches we will use for **Step 1** (modeling the decision to apply for disability benefits). This sub-section includes a discussion of the literature related to application for disability benefits, dynamic modeling, and microsimulation. We will describe three modeling approaches: (1) Hazard Modeling, which is a reduced form approach; (2) the Option Value Modeling approach, which is a structural modeling approach; and (3) Dynamic Programming Modeling, which is also a structural modeling approach.

After describing these three approaches, we will explore some dynamic dimensions of the disability benefit structure: (1) how eligibility for SSDI and SSI benefits changes over time; (2) how benefits change over time; and (3) how an individual's future wage earnings influence his or her future benefits. These dimensions will need to be incorporated into our dynamic programming modeling. We also discuss certain key factors which influence whether or not a person applies for disability benefits.

Step 2 is the modeling of the SSA decision to allow or deny benefits. We summarize how we will do this, and include sections on the initial determination process and on modeling the appeals process. A discussion of **Step 3** (applying the models to predict the future impact on program size) is next. This is our last step to meeting this research objective. Although it may appear to be the most straightforward, it can only be accomplished with the results of Steps 1 and 2 that precede it. Finally, we end this section with a discussion of the **data elements** required for this modeling process.

STEP 1: MODELING THE DECISION TO APPLY FOR DISABILITY BENEFITS

While there is little doubt that future applications for disability insurance will be driven by demographics and the prevalence of disability in the population, it is clear that economic conditions and policies with respect to accommodation and disability determination are also likely to be important. The influences of SSDI and SSI programs on the decisions of men and women with serious health conditions to leave the labor force has been in dispute for over a decade. The consensus is that changes in morbidity, economic variables (e.g., general economic conditions and willingness of employers to accommodate), and disability policy variables (such as ease of qualifying for benefits and the relative benefit level), all play important roles in explaining variations in disability application rates over time.

Burkhauser, Butler, and Kim (1995) studied data from the 1978 Survey of Disability and Work. Using a discrete-state, continuous-time hazard model they found that the expected replacement rate of disability benefits and accommodation by an employer affect the risk of job exit following the point at which a health condition first begins to limit work. Burkhauser, Butler, Kim, and Weathers (1997) looked at how policy variables affect the speed at which workers apply for disability benefits following the point at which a health condition first begins to limit work. They extended the Manski and Lerman (1977) correction of a choice-based sample to a likelihood function with semiparametric unmeasured heterogeneity. In addition, they greatly improved the measures of values of wage earnings alone, and interacted with a dummy variable for having a health condition that limits the kind or amount of work one can do. They then contrasted these results for key policy variables.

First, they separated the replacement rate into a time-varying variable equal to the actual Social Security benefit a worker would receive in each period if she/he were admitted onto the disability benefit rolls. Second, they predicted wage earnings in each period using autoregression—a constant and four lagged values of wage earnings alone and interacted with a dummy variable for having a health condition that limits the kind or amount of work one can do. More importantly, the paper then contrasts these results for the same model but using data from Wave 1 (1992) of the Health and Retirement Study. Policy variables were shown to importantly influence behavior using these two data sets fielded 14 years apart.

Both of these papers were attempts to look at the transition out of the labor force and onto the disability rolls by people with disabilities. They are subject to the criticism that they did not contain an independent measure of a worker's probability of acceptance onto the benefit program. We propose to extend this line of reduced-form estimation using our more sophisticated measures of probability of acceptance onto the disability program discussed below, and by including a broader set of human capital-based explanatory variables. But we will also use structural models. Neither of these papers nor our proposed extension of them makes use of the more sophisticated dynamic models that have been used in the retirement literature. We propose to do so here and then compare the powers of our reduced-form model with that of our structural models in predicting future growth in program population and costs.

The two major innovations in the use of dynamic decision-making rules are those based on an option value approach (see especially Stock and Wise, 1990; and Lumsdaine, Stock and Wise, 1990), and those based on dynamic programming (see especially Rust, 1989; Berkovec and Stern, 1991; Daula and Moffitt, 1995). While these models have primarily been used in the retirement literature, we believe they have potential application to the decision to apply for disability benefits as well. The application of the models to the SSDI population is developed here. The application of similar models to SSI claimants will be discussed later.

Both the option value approach and the dynamic programming approach allow individuals to recalculate their optimal behavior each time period, using new information about the process and their best prediction about the future. What makes these models particularly interesting is that they do not assume that workers know either their future wage rates or their retirement benefit flows with certainty. Thus, they are much more realistic than earlier life-cycle models. (For examples of perfect certainty retirement models, see, for instance, Burbridge and Robb, 1980; and Fields and Mitchell, 1984.).

Stock and Wise (1990) develop and test the option value model for workers from a single firm. While they do not estimate a dynamic programming model, they explain the difference between the two approaches—the option value model evaluates the future as the maximum of the expected values of utility, whereas the dynamic programming model evaluates the theoretically more appropriate expected value of the maximum, which is necessarily larger. Lumsdaine, Stock, and Wise (1990, 1992) compare the dynamic programming, option value, and reduced form probit models of retirement from a given firm. They find that the structural models are about equally accurate in predicting the effects of a window plan. Both models are more accurate than probit, but the option value model is easier to estimate. Neither predicts actual retirements at age 65, an age which is more popular than economic factors can explain.

Dynamic programming models have been used for other labor supply issues. Daula and Moffitt (1995) extended their earlier work (Daula and Moffitt 1991) by considering more than two periods in a dynamic programming model of Army reenlistment. They propose some simplification of the estimation and emphasize the ease of estimation of their model. Ausink and Wise (1993) study the retention of Air Force pilots. They apply the models of Lumsdaine, Stock, and Wise (1990), comparing dynamic programming and option value models. This time the option value model more accurately predicts behavior. These new approaches will form the basis for modeling the decision to apply for disability benefits in this proposal.

Douglas Wolf is a consultant to the DES project who has been involved in the development and application of microsimulation techniques for many years. The *methodological* goals of his projects are to place microsimulation on more rigorous grounds through the specification and estimation of multidimensional models that recognize interdependencies of both the systematic and the stochastic elements of the models, and to quantify the contribution of several sources of stochastic uncertainty to overall uncertainty of simulated outcomes. The techniques used to quantify uncertainty include variants of multiple imputation (Freedman and Wolf, 1995), randomization over ex post distributions of model parameters, and resampling techniques such as the bootstrap (see Wolf, et al., 1995).

The *substantive* goals of his project are to develop models of the joint evolution of kinship networks, economic well-being, and disability within the older population, and to produce disaggregated projections of the future older population of the United States using the model. McNally and Wolf (1996) illustrate the use of multiple imputation to construct a sample of individuals that represents the initial conditions of the dynamic model, by combining observations from two independent but overlapping sampling frames. Applications of the randomized-parameter and bootstrap techniques to quantify uncertainty can be found in Wolf and Laditka (1996) and Wolf (1997a, 1997b). We plan to develop a set of alternative parameters for key policy measures using our alternative models and data from the DES for inclusion in microsimulation models developed in collaboration with Wolf.

Hazard Model

Burkhauser, Butler, Kim, and Weathers (1997) used a reduced-form hazard model to estimate the time to application following the point at which a health condition first begins to limit one's ability to work. We will extend that work in this project by using the more independent measure of probability of receiving disability benefits. We will also add additional explanatory variables as discussed below. We will use this more sophisticated but nonetheless reduced-form model as a basis for comparison with the structural models we will develop. In addition to the policy variables, this model will include variables that are based on human capital theory, some of which will also be discussed in an earnings function context in the structural models. These will be the explanatory variables in the hazard model:

- **Expected Earnings.** We expect to obtain Social Security earnings histories from the Social Security administrative records. We will use the earnings histories to estimate future earnings for the individual using an autoregression. Higher earnings increase the opportunity cost of applying for disability benefits. We expect that individuals with higher expected earnings are more likely to continue work and postpone application for disability benefits.

- **Primary Insurance Amount (PIA).** We will use the Social Security Earnings Records to calculate Social Security benefits for our sample. We expect that individuals with higher expected PIAs will place a higher value on the application option and are more likely to apply sooner than individuals with lower expected PIAs.

- **State-Time Allowance Rate.** State-time allowance rates are defined as the number of individuals accepted into the disability benefits program divided by the total number of applicants. The state-time allowance rate is the source of exogenous variation that we use to measure the individual probability of being accepted into the program. The higher the probability of acceptance, the sooner one will apply. These data are available for 1974 through 1993. It will be necessary to request SSA to update these data for this study.

- **State-Time Unemployment Rate.** The State Unemployment Rates from 1970 through 1993 are included in a public use dataset. We will update these values. Poor business cycle conditions, as measured by the state unemployment rate, are likely to reduce employment opportunities following the onset of a work-limiting condition. We expect high unemployment rates to lower the expected value of the continued work, so that individuals will be more likely to apply.

- **Employer Accommodation.** Accommodation can affect the relative benefits of work compared to SSDI application in several ways. The health limitation should be mitigated to some degree by employer accommodation. Therefore, workers who are offered accommodation are more likely to utilize their remaining stock of human capital and be willing to accumulate additional human capital. Furthermore, employee accommodation might take the form of retraining the worker for a new job. As a result, the benefits of continued work are likely to be higher for those who were offered accommodation and to make them less likely to apply.

- **Job Characteristics.** Direct measures of job characteristics prior to the onset of a work-limiting condition will be incorporated into the retrospective questionnaire. These measures have been categorized and used in recent work by Daly and Bound (1997). We will use these measures to capture the effect of job characteristics on the length of time to SSDI application in our models. In addition, we will use the broader occupation groups to define white collar and blue collar workers. Individuals who perform blue collar work or work that is more physically demanding compared to other types of work are likely to find that their human capital depreciates at a faster rate than does that of white collar workers. Hence, one would expect that new investment in human capital following a health limitation would be less attractive for blue collar workers and, therefore, these workers would apply for SSDI sooner.

- **Industry.** The individual's industry at the time of onset of a work limitation will be available in the DES questionnaire. Industry is classified into 13 categories based on original United States Census codes. Differences across industries in prospective growth and use of technology should affect the path to application. We expect those workers in low growth or high technological change industries will apply for disability benefits sooner.

- **Union Status at Onset.** Retrospective questions regarding union status on previous jobs as well as union status on the current job will be available. We expect that individuals who are members of a union are more likely to have contracts that protect their jobs and, thus, to postpone disability application.

- **Employer Size at Onset.** Questions regarding the firm size at the plant as well as at all locations will also be asked. To the extent that larger employers are more likely to have "internal labor markets" or implicit contracts that pay a worker less than their marginal product for some period and then more than their marginal product at some future period, the size of the firm for whom the employee works may affect the possible path chosen. Given that these future earnings above an individual's marginal product require continued work, one would expect that the work option would be relatively more valuable and that individuals who work in larger firms will be more likely to postpone disability application.

- **Health Conditions at Onset.** Our data will have excellent information on health conditions. We will use several different definitions for health conditions. We expect differences in the rate of depreciation in human capital resulting from different types of work-limiting conditions. For example, **musculoskeletal conditions** tend to be degenerative and the depreciation of human capital among workers with such a condition should occur at a relatively slower rate than in individuals who have **cardiovascular conditions** which are relatively more sudden and debilitating. Westat's medical consultants to this project will inform the modeling of the health condition variables.

- **Education.** Individuals with higher levels of education are assumed to be better able to learn new skills and adapt to changes in their ability to perform paid work. Therefore, we expect that individuals with higher levels of education will be more likely to postpone disability application.

- **Tenure on the Job.** Individuals with greater tenure on a job should have made greater investments in specific human capital on that job and should be more willing to adapt to work-limiting conditions on that job. We expect that these individuals will find the benefits of continuing to work greater than will individuals with lower levels of tenure, and thus will postpone disability application.

- **Age at Onset.** Individuals who are older invest less in human capital because they have a shorter time horizon to reap the benefits of the additional investment. Similarly, individuals who experience a work-limiting condition at older ages should have less of an incentive to invest in the human capital necessary to overcome the work-limiting condition due to the shorter time horizon available for realizing the benefits of such an investment and will choose to apply sooner for disability benefits.

- **Marital Status at Onset.** Married men tend to work more hours than single men; therefore, similar rates of investment in human capital over more hours of work would suggest steeper age earnings profiles, as well as higher levels of earnings, for married men compared to single men. We expect that married men would therefore postpone application for disability benefits. Based on our own analysis of the National Health Interview Survey - Disability Supplement (discussed below), we also expect those who do not live alone to be less likely to apply for benefits due to the physical and psychological assistance another person in the household can provide.

While this description relies on an economic approach, we will also attempt to expand these models to include the other influences on disability that were indicated in figure 1. Thus, we have described here that analysis which we are confident we can deliver. We have not tried to describe our hopes for further and deeper analysis using non-economic variables that may be much harder to quantify.

Structural Models

Here we present overviews of the option value and the dynamic programming models that we propose to use to investigate the timing of application. Each builds on the work of Lumsdaine, Stock, and Wise (1990) but are significantly adapted to the unique issues related to application for disability benefits. The reader who is interested in the specific details of the models should go to Appendix A for a more complete mathematical exposition.

Application for disability benefits occurs when applying for benefits yields a higher utility now than at any future time; otherwise application is delayed at least one period. This outcome depends nonstochastically on the factors discussed below and on stochastic disturbance terms. To calculate the utility gains from applying for disability benefits at different points in the lifecycle, one must first specify the choices to be modeled as well as the utility functions and the distributions of the stochastic elements.

The choice set in any year consists of continuing to work or applying for disability benefits. The outcomes of applying for benefits are either acceptance or rejection. Rejected applicants must either return to work (likely at a lower expected wage) or not work. Rejected applicants who return to work might return to the previous job or move to another job. Since the choice of application is available in any year, application can be postponed into the future, or may never occur. In this latter case it is assumed that retirement occurs at age 65 when eligibility for SSDI ends. Reapplication is also possible, and rejected applicants each year consider whether to reapply in subsequent years.

This model focuses on the decision to “retire” prior to age 65 for those with permanent health conditions that affect their ability to work. Because all beneficiaries automatically shift to the Social Security retirement program at age 65 or are assumed to retire at age 65 if not on SSDI, we abstract from issues of how long such workers will live past age 65. Because our data will be retrospective, everyone in the dataset is alive so it is not possible to study earlier mortality. Therefore, in the model we assume all persons live to at least age 65. However, for purposes of developing cost predictions associated with alternative definitions of disability or decision-making processes, this assumption can be relaxed.

The dynamic programming version of this model uses most of the same equations, but with different probability computations at the end (see Appendix A). The calculations involve the expected maxima of utility over times of application, beginning at the end (retirement at age 65) and working backward. Independent or possibly random effects are assumed by Berkovec and Stern (1991) and Daula and Moffitt (1995). Keane and Wolpin (1994) show how to estimate models with more complex error structures, but the models we propose do not require simulation techniques to solve.

Using these methods with the DES data, we will estimate coefficients in the equation determining net benefits of work and SSDI application. This will be done using various measured and estimated inputs, including the shifts of expected earnings when a disability occurs. Based on these results, we can simulate the effect of various policy changes on the future work paths and SSDI application paths of workers following a disability, and we can see how greater accommodation on the job, the reduction of SSDI benefits, or the reduction in acceptance rates can change these outcomes.

Dynamic Dimensions of the Disability Benefit Structure

Lumsdaine, Stock, and Wise (1990) developed their model to show how the structure of a retirement program in a single firm influences the decision of workers to retire from that firm. In our models we propose to show how the structure of Social Security disability benefits influences people with disabilities to apply for benefits. To see how we will operationalize these models, we need to discuss three critical dimensions of the disability benefit structure: (1) how eligibility for SSDI and SSI benefits changes over time; (2) how benefits change over time; and (3) how an individual's future wage earnings influence his or her future benefits.

(1) How eligibility for SSDI and SSI benefits changes over time

What makes the decision to apply for disability benefits more complex than the decision to apply for retirement benefits is the possibility of benefits being denied. For instance, to receive Social Security retirement benefits one must have a sufficient number of quarters of Social Security coverage and be age 62 to be eligible to receive early benefits, or age 65 to be eligible to receive normal retirement benefits. To be eligible for SSDI benefits, in addition to a quarters-of-coverage requirement, the state-administered Social Security "substantial gainful activity" (SGA) test must be passed. While the SGA eligibility rules are the same across states, over the last two decades the denial rates have varied dramatically over time and across states. Hence, in addition to using the socio-economic characteristics of individuals who have applied for benefits in an estimation equation to assign acceptance probabilities to our respondents, as has been done in previous work, we will use an exogenous acceptance probability measure based on program administrative behavior.

Table 8 shows the large degree of variation in the acceptance rate (allowances per application) across states and time. Since we are interested in state variation, we report the mean of the state rates unweighted by state populations. Aarts, Burkhauser, and de Jong (1996) argue that the “demand” for disability transfer recipients as voiced via both program rules and their administration may be as important as the “supply” of potential beneficiaries in explaining differences in disability transfer populations across countries. As Table 8 shows, mean state allowance rates varied from a 30.7 percent low in 1982 (a period of major retrenchment in the system) to a high of 50.1 percent during a period of great program expansion in 1975. Furthermore, the rapid increase in program population in the early 1990s is coincident with acceptance rates that were nearly as high as those of the mid-1970s. Variation across states is even larger. Probabilities of acceptance vary across states from lows of around 20 percent to highs of over 65 percent over this period.

These data measure the SSDI administration climate in the states and introduce exogenous variations into the SSDI application decision caused by differences in state administered application of SSDI eligibility rules. By using these data, we avoid the endogeneity of individual decisions and, given that Parsons (1991b) shows that individuals make reasonable guesses about the probability of approval, this statewide measure should proxy well for individual perceptions. Identifying the model using nonlinearities and functional form assumptions is possible, but given the potential complaints about any set of assumptions, we would rather use the differential allowance of states using these geocoded variables. They are arguably exogenous, their exogeneity can be tested, and they provide statistical variation that usually identifies models more robustly in selection bias and simultaneous equation systems.

The probability of acceptance can be estimated using clearly exogenous variables such as gender, race, and age. We can easily add such variables to the state and year data to improve the prediction of the probability of acceptance. Whether the probability of acceptance should be lagged is a theoretical question, since such a change is unlikely to affect the model significantly. We expect that most applicants are well informed of broad changes in behavior by doctors, attorneys, and examiners. Using lagged values would present no major problems apart from a small loss of sample size. These data are available at the state level from 1974 through 1993. We expect that the Social Security Administration will make them available to us through 1998.

(2) How Benefits Change Over Time

The value of SSDI benefits that workers may receive at any time is based on their Social Security covered earnings to that time and the formula for calculating benefits. The formula for calculating benefits has changed over time, and the potential benefit each worker could have received at each year following the onset of a disability would be quite difficult to estimate without detailed information on that worker's Social Security earnings record. (See U.S. DHHS 1995 for the historical formulae for SSDI benefit calculations.) This analysis is predicated on the assumption that the Social Security earnings records will be available for those who grant Westat permission to use them for the DES. SSI benefits are set by statute, and are identical for all beneficiaries. The historical time series for SSI benefits will indicate how benefits have changed over time. These can be adjusted for changes in the cost of living, which will introduce some additional variation across years.

Using these data, we will calculate a present discounted value of disability benefits for all workers for all years based on the Social Security rules in place in each year, the workers' earnings record to that point, and a projection of their future wage earnings. For workers who postpone application, we will recalculate Social Security values based on their next year of actual earnings, actual changes in Social Security rules in that year, their projected future wage earnings, and so on for all additional years until application. Burkhauser, Butler, Kim, and Weathers (1997) provide an example of how administrative records can be used for this purpose.

(3) Influence of Future Wage Earnings on Future Benefits

Future SSDI benefits are affected by future wage earnings. There are several methods of predicting future wage earnings. Blau (1994) uses a standard labor economic model, with individual characteristics such as age, experience, and tenure, in a fixed effects model. We will follow Burkhauser, Butler, Kim, and Weathers (1997) and borrow a method from macroeconomic models to predict labor income using autoregression. The models include a constant and four lagged values of labor income, first alone and then interacted with a dummy variable that indicates that the worker had a limitation on the kind or amount of work he/she could do. Their prediction equation produced 88 to 92 percent accuracy

Table 8. Descriptive Statistics of the Distribution of State Allowance Rates Over Time[a]

Year	Mean of State Allowance Rates[b]	Standard Deviation	Minimum State Allowance Rate	Maximum State Allowance Rate
1974	0.491	0.063	0.379	0.634
1975	0.501	0.063	0.381	0.656
1976	0.476	0.056	0.355	0.582
1977	0.451	0.056	0.333	0.598
1978	0.395	0.067	0.230	0.524
1979	0.361	0.067	0.225	0.548
1980	0.341	0.055	0.227	0.473
1981	0.315	0.047	0.232	0.421
1982	0.307	0.047	0.211	0.409
1983	0.336	0.059	0.235	0.465
1984	0.358	0.063	0.232	0.518
1985	0.373	0.067	0.267	0.531
1986	0.396	0.067	0.266	0.525
1987	0.371	0.067	0.196	0.499
1988	0.367	0.068	0.206	0.527
1989	0.373	0.067	0.218	0.519
1990	0.392	0.068	0.225	0.505
1991	0.411	0.071	0.250	0.551
1992	0.421	0.070	0.248	0.564
1993	0.385	0.063	0.260	0.490

[a] From 1974 to 1979 these data were provided by Donald Parsons for all states except Alaska, the District of Columbia, and Hawaii. From 1980 to 1993, SSDI allowance rates were provided by the State Disability Determination Services for the 50 states and the District of Columbia.

[b] Defined as the total number of SSDI allowances divided by total number of SSDI applications.

NOTE: These are the mean of state values, not the national mean. That is, each state regardless of populate population is given the same weight in the distribution.

Source: Public Use Data File on Social Security Disability Insurance and Supplemental Security Income Applicants and Beneficiaries prepared by Lewin-VHI, Inc.

values. Human capital variables which do not change are included via the lagged earnings. Individual specific earnings effects are also accounted for that way. Variables that change over time also may be included, however. These include accommodation, and state-specific, time-specific state-time unemployment rates.

We plan several extensions of this equation. We will include "accommodation" (whether an employer accommodated the worker at the time his or her health condition first began to affect his or her work, or whether a worker began a new job in which he or she was accommodated). We will then interact accommodation with socio-economic characteristics of the worker, such as age, past labor earnings, education and gender. These interactions allow us to vary the effect of accommodation by demographic characteristic and to estimate the gains from accommodation across socio-economic groups. Burkhauser and Daly (1997) estimate from a sample of 1,209 disabled workers that 22.2 percent are accommodated, with large variation by demographic characteristics.

Some persons in our sample may work after they have been denied disability benefits. We will include a dummy variable to account for that event. We will try a tobit specification to take account of the fact that labor earnings may be zero. We prefer to use an equation based on a feasible information set, to mimic individual decisionmaking and to allow simulations based on changes in accommodation. We propose to use only tobit models as possible prediction equations for sources of income, which are often zero. Again, selection and specification could be quite complex for these models, but for the purpose of prediction all of that may be ignored.

Work Following Disability Insurance Application or Acceptance

To receive disability benefits one must have a physical or mental impairment that has prohibited work for five months and will make it unlikely that work can be performed for at least one year. The work test is based on severity. If the impairment is severe enough, one should not be able to perform any substantial gainful activity (SGA). If one earns more than $700 per month ($500 per month before July 1999), this is prima facie evidence of SGA. Under certain circumstances, earnings above $700 are permitted for short trial periods, but few people who apply for disability benefits work during the application process, or while receiving benefits.

Persons may return to work if they are denied disability benefits. Those who do so apparently earn, on average, significantly lower wages than they earned before they applied for benefits (Bound, 1989). We propose to try two alternative assumptions in the estimation. First we assume that the decision to apply for disability benefits is a decision to leave the labor force permanently. Because the appeals process can extend for a year or more for those who take advantage of all avenues of appeal (Parsons, 1991a), during which time applicants cannot work without diminishing their chance of acceptance, this is a reasonable first approximation. Under our second assumption, applicants have some probability of reentry to the work force. We will estimate that probability and expected labor earnings based on information from previous research, such as that of Bound (1989).

Some persons turned down for disability benefits probably can return to work. We will include four possible outcomes: no application, application and acceptance, application and rejection with return to work, and application with rejection and no return to work. We assume the utility function of earnings is the same before and after application, although earnings are probably smaller because of the disability and time away from work, as the autoregression will show. This complicates the model in that four possible paths must now be investigated in the estimation, but at most there are two disturbances.

Other Variables

The DES will include information about the following individual factors (see figure 1); health status, functional limitations, and personal characteristics such as education, training, work history, etc. Environmental factors will also be included; such as social supports, access to community, work environment, access to health, education and vocational services. All these factors play important roles in determining disability status. When information about the onset of disability is combined with social security earnings records we will have pre- and post-impairment work histories for all individuals impaired at the time of the interview. The exogenous variables we propose to use include age, race, gender, education, marital status, job characteristics, industry, occupation, and state of residence. Table 9 lists the variables that will be available in the DES and gives the location of the corresponding question(s) in the comprehensive survey instrument (CSI).[7] A method of linking SSA records to these data will be developed to satisfy confidentiality concerns with respect to administrative records.

At this point the behavioral models include predominantly economic variables. Complex dynamic models at this stage do not permit the broad-ranging variables found in labor supply equations, for example. Still, we think we have captured the essential features of the structural decision process facing the potential applicant for SSDI; the forward-looking element which is so vital in individual decisionmaking.

---

[7] This list refers to the CSI version submitted to OMB in June 1999. It may not correspond exactly with the final instrumentation.

Because this study will have high quality data on the medical conditions and functional limitations behind the specific type of work-limiting condition, we will be able to see how different categories of health conditions produce different outcomes. We assume these health conditions and other individual factors are exogenous factors affecting wages and are thus critical in the decision to apply for disability benefits. But they may also have other, more direct effects. Thus, some impairments are so immediate and so severe (a major stroke, for example) that they pre-empt the careful rational choice process underlying our models. Similarly, for environmental factors. Sometimes, an application for SSDI or SSI may be precipitated by a policy change in an unrelated income maintenance program, that suddenly removes the income support from a disabled individual. Our period-by-period rational calculation models will have to be modified for these influences as well.

In Burkhauser, Butler, Kim, and Weathers (1997), we did not predict nonwage income, but here we propose to do so, using autoregressions augmented by other explanatory variables, and possibly tobit models if they improve the predictions. Since nonwage income consists of returns on assets and payments from programs with predictable flows, we expect accurate predictions. Several sources of income can be used to predict each other in a vector autoregression. The vector autoregression would relate all sources of income (labor income, transfers, and other income), to each other and to an indicator for a condition limiting the kind or amount of work one can do.

We doubt other variables would help much, given the results of labor income predictions, and we expect relatively high predictive power. We do not seek structural models at this point, just good predictions. We believe that individuals use such methods implicitly in their own decisions. We will be able to observe long time series of wage income for many people with the Social Security earnings data. Lagged dependent variables will implicitly account for fixed effects.

Accommodation by employers can increase the length of time during which an employee stays on a job and does not apply for SSDI (Burkhauser, Butler, Kim, and Weathers, 1997). Employer accommodation will be measured in the DES. We propose to control for this variable in the dynamic models, and we expect it to be an important factor. Rather than representing the effect as a constant shift in the wage, however, we propose to show the effect of accommodation on SSDI application in a random coefficients model, interacting accommodation with various health conditions, age and other factors. In this way, the effects of accommodation can be different, and the benefit of policies that increase accommodation can be investigated. All of these measurements and predictions will be inserted into the dynamic programming and option value models to analyze people's decisions to apply for SSDI.

Table 9. Questions Contained in the CSI Relevant to Disability Modeling

Variable	**CSI Questions**
Earnings	WS-85
Employment History	WS-62 to WS-65
Wealth	ER-117 to ER-169
Income	ER-1 to ER-116
Primary Insurance Amount	ER-34 to ER-51
Employer Accommodation	WS-114
Job Characteristics	WS-80, WS-100, WS-104
Industry	WS-78 to WS-79
Union Status	WS-84
Employer Size	WS-81 to WS-83
Educational Attainment	HE-10 to HE-11
Job Tenure	WS-64 to WS-65
Hours Worked	WS-67 to WS-69
Age, Race, Gender	HE-6 to HE-9
Marital Status	HE-12
Medical Events and Conditions	MH section
Welfare History	ER-34 to ER-110
State of Residence	HE-13
Discrimination	WS-128 to WS-131
Rent or Mortgage Payments	ER-122 to WS-181
Program Knowledge	ER-67, PK sect.
Work Conditions	WS-106
Worker Preferences	WS-107
Job Satisfaction	WS-108
Social Companionship, Assistance, Support	SC-1 to SC-2
Living Arrangements	HE-1 to HE-5, SC-20
Health Insurance Coverage	AS-1 to AS-38
Training or Employment Services	AS-41 to AS-51
Specialized Services (i.e., physical therapy)	AS-52 to AS-63
Transportation Availability	SC-14 to SC-19

In the absence of additional waves of interview data, we will use a wage equation based on information prior to SSDI application to predict labor earnings in each year in which a worker may decide to apply for disability benefits. We do so by allowing labor earnings to vary based on past earnings, accommodation, disability, and accommodation interacted with socio-economic variables. This interaction of accommodation with other variables could prove to be difficult to model, and we cannot guarantee that all of the proposed parameters in our dynamic models will converge or be estimated successfully. We have proposed what we think can be estimated based on our experience in estimation of nonlinear models.

Finally, we must predict household income for persons who are denied disability benefits and do not return to work. We assume this would consist of income from alternative government transfer programs (e.g., welfare, food stamps, etc.) and support from other family members. Information on state level general assistance and food stamps over the period are available, but we will have no information on the income of other family members prior to the survey date.

MODELING SSI APPLICATION

While the application process is the same for SSDI and SSI, the qualifying characteristics are quite different, and thus, the analysis of applications will also differ. Table 7 listed the variables used by various researchers to estimate applications for disability benefits. Most of the variables listed are subject to change over time and should be measured from the onset of the disability—e.g., primary insurance amount, state-time allowance rate, state-time unemployment rate, job characteristics, industry, union status, employer size, education, tenure on the job, marital status, medical events, etc.

To date no one has specifically attempted to study application for SSI using the type of models discussed above. To do so would require many of the same variables used for SSDI claimants. But because a large share of SSI applicants may not be in the labor force at the onset of their disability, other variables may be more important. For instance, participation in and income from other welfare programs (e.g., food stamps, general assistance, AFDC, TANF, etc.) may affect the SSI application process. Other income from child support or alimony would also be important, as would family structure beyond marital status–e.g., the presence of dependent children or adults, age structure of children, or extended families, all indicators of social supports.

Of course, each of these variables should be measured at onset of disability, which is a practical impossibility with a retrospective study design. All variables should be based on questions that clearly distinguish between the year before onset and the year of onset, because some disabilities immediately reduce income and potentially change family structure and job characteristics. The problem is endogeneity, (since some are determined jointly with application for SSDI or SSI). Thus, the data should generally cover the year before onset and all years since onset.

Like our models of SSDI application, the decision to apply for SSI can be estimated with reduced form logit, probit, or hazard models, and we plan to do so. Dynamic programming and option value models assume that the decision to apply for program benefits depends on the value (utility) the person places on income from work and from disability program benefits. Burkhauser, et al. (1997) limited their analysis to those who experienced onset of a disability while working. The model focuses on the fact that once a person has a disability which creates a nonzero probability of being accepted onto the SSDI program, a valid comparison can then be made between her or his ceasing to work and applying for SSDI or continuing to work. This may not be the case for persons who are potentially eligible for SSI. We consider children and adults separately.

Beneficiaries receiving SSI since childhood may have no work history, and hence the dynamic programming or option value models discussed above don’t capture the factors influencing their behavior. While some variables in those models should also predict application for SSI, they would represent, at most, utility differences unrelated to SSI or to SSI income, and hence not representative of the sort of comparison envisioned in our structural model. Consequently, while we could estimate an equation predicting SSI application for children or other non-workers, it would not be one comparable to the SSDI models we have discussed above. A salient issue with respect to childhood onset beneficiaries is how many such individuals find jobs and leave the rolls as young adults. This question can be answered with a reduced form hazard, probit, or logit model, but lacking a good measure of the potential wage which the young adults could earn, dynamic programming and option values cannot be estimated.

Some adults eligible for SSI also may have no work history. As with children, we are able to predict the decision to apply for SSI benefits, but not in a manner comparable to our structural models. Adults with a limited work history are a different matter, however. (Persons with a qualifying disability but with a work history insufficient for SSDI eligibility.) Such workers at least have some dynamic history of work and earnings, and so they theoretically face the same kind of decision that is relevant for SSDI applicants, although the options may be less palatable. It is likely that their potential wages and their income benefit is lower, relative to those of persons considering SSDI, but the estimation of the model will determine whether income has the same utility for both groups, or whether the processes are fundamentally different. Measuring wages of irregular workers using administrative records can be difficult however, so the outcome is unclear.

In Table 10 we provide a list of variables we will include in our SSI participation models both for persons with no work history and persons with a limited but nonzero work history. Education, age, sex, and race would be included as is typical in models of program participation, for the usual reasons: utility differences, cultural differences, or discrimination. Other types of income received at onset, or for periods following onset, (welfare programs, alimony, child support, income from assets) would affect the willingness to apply for SSI. Even given that the sum is small, the source might matter either as an interaction of programs or as a signal of utility differences. Wealth is unlikely to be a factor, because it is restricted by the rules governing eligibility. Family structure matters, such as the presence of a spouse, children, or other dependents. The medical condition which causes the disability can definitely affect applicant behavior. In addition to an appropriately coded medical condition variable, the self-report of health might have an additional effect as an indicator of the degree of severity of the condition, or of the attitude of the potential applicant. Finally, some states have encouraged potential applicants for SSI to apply. If a measure of state recruiting can be found, we will match it to state of residence of the respondents.

As we have discussed above, we will use reduced form hazard, probit, or logit functional forms to model application for SSDI. We will also apply more elaborate structural modeling which allow us to explicitly compare utility in different states (applying or not applying). We will apply the same reduced form models to SSI applicants, but in many cases it will not be feasible to use structural models, because there is no realistic choice of streams of income to compare.

Table 10. SSI Application

Variables	Work History	No Work History
Expected Earnings	X	
Expected Employment	X	X
Earnings in Preceding Year	X	
Employment in Preceding Year	X	X
Wealth at Onset	X	X
Income at Onset	X	X
Primary Insurance Amount at Onset and Thereafter	X	X
State-time Allowance Rate at Onset and Thereafter	X	X
State-time Unemployment Rate at Onset and Thereafter	X	X
Employer Accommodation at Onset	X	
Job Characteristics at Onset	X	
Industry at Onset	X	
Union Status at Onset	X	
Employer Size at Onset	X	
Education at Onset	X	X
Tenure on job at Onset and Thereafter	X	
Hours Worked at Onset	X	
Age, Race, Gender	X	X
Marital Status	X	X
Medical Events and Conditions	X	X
Activity Limitations (ADLs, IADLs, etc.)	X	X
Work Experience and Demands	X	
Vocational Skills	X	
Vocational Rehabilitation	X	X
Work-Incentive Provisions	X	X
Reasons for Working	X	
Welfare History at Onset	X	X
State of Residence	X	X

STEP 2: MODELING DISABILITY DETERMINATION

Initial Disability Determination

SSA determines the eligibility of applicants for disability benefits through a five-step screening process. Each step is based on different information. Lahiri, et al. (1995) modeled the disability determination as a sequential logit model incorporating each of the disability determination steps. The purpose of their multistage model was to capture the structure of the process so that proposed policy changes could be simulated, and to estimate the pool of eligibles in the general population. Such a model is useful in forecasting the future pool of eligibles and thus the trend in disability caseloads. With the data from the DES, including the simulated DDS determinations of eligibility, we will be able to model these processes very effectively.

Application, appeals, and hearings before an ALJ can be modeled in a reduced form which ignores the dynamic structural, utility aspects of the process, or they can be modeled in a structural form which considers the parameters of the utility function. The existing literature uses reduced forms, so we emphasize them here. Benítez-Silva, et al. (1998), find that estimating the appeal process separately from the determination process is important for understanding the extent and source of Type I and Type II errors in the administrative process. Lahiri, et al. (1995), discussed whether a structural model predicted allowances and denials more accurately than a reduced-form model. They argued that a structural model is preferred because it is difficult to identify the correct effect of variables on the final eligibility probability unless they are introduced at the appropriate stages. We will explore the merits of maintaining a distinction between these stages of the process in the models used to predict future caseloads.

Estimation of Appeals and Decisions of Administrative Law Judges

The object is to estimate the probability of application for disability benefits and the probability of appealing a negative determination. We also need to estimate the probability that an ALJ will reverse a denial. Because accommodation on the job is a major factor affecting the application for SSDI, it is also important to model the probability of being accommodated. Our first approach to this is to estimate a sequence of independent logit models for as many of these stages as possible: accommodation, application, appeal, and ALJ decision. Econometric identification of the stages is achieved by the assumption of independence of the disturbances. Independence is assumed precisely because it is so difficult to identify the various stages of the process by finding variables which affect one stage but not another. Even if the disturbances are independent, however, this approach is not necessarily the best one to predict the various stages of the process.

If accommodation by the employer of one's physical or mental disability is exogenous to the type of work performed and other outcomes of the labor market, then a dummy variable can be inserted into the model for application for disability benefits. Appeals would not be affected, because people appealing negative decisions are virtually never currently working. Accommodation may be endogenous, however, as a result of unobserved productivity differentials (unobserved by the researcher, that is); then an equation is required to adjust for the endogeneity. For this purpose, data describing the job and other personal conditions (e.g., marital status, wages) at the time of onset of the condition are vital. This is another reason to collect retrospective data. The decision to accommodate likely depends on the benefits to the employer of continued employment by the affected employee and, particularly since the passage of the Americans with Disabilities Act, on the cost of accommodation. Thus, the variables needed to build such a model are the costs and benefits of accommodation, or proxies for them.

The purpose of estimation can be to specify and test structural models or to make better predictions. In a dynamic model of choice, it can be more important to have good predictions than to have good descriptions of the actual process used by ALJs, for example. Considering that, there may be better approaches than logits, even given the assumptions made about the disturbances. Fundamentally, the best predictions are the most accurate ones, and the method of estimation which maximizes the explained variance is linear regression. Thus, linear probability models, for all of their problems (e.g., predictions may be out of range) usually make the most accurate predictions. Thus, substituting linear probability models for more conventional limited dependent variable models may be worth considering in the interest of increasing the explanatory power of the model.

A second approach is to exploit categories. Consider such variables as sex, race, and type of condition. If these can be considered exogenous to the process, then the data can be segmented into subsamples for which separate models are estimated. This approach can increase considerably the accuracy of estimation, at the cost of having multiple equations to interpret. More categories reduce the usefulness of individual coefficients for analysis but increase the usefulness of the overall model.

Option value models and dynamic programming models of the utility of applying for disability benefits or appealing a negative decision are presented above. These models are an alternative to the reduced form models, which are confined to predicting outcomes. However, the only dimension on which the models can all be compared is predictions of outcomes, and we propose to do that, just as Lumsdaine, Stock, and Wise (1990) did with retirement models.

STEP 3. APPLYING THE MODELS TO PREDICTING FUTURE PROGRAM SIZE

The models of the application and award processes estimate the propensity of an individual with a perceived disability to apply for benefits (step 1) and the likelihood that an individual will be awarded benefits (step 2). To estimate future trends, we also must project into the future the number and characteristics of people who will be eligible for disability benefits. As already mentioned, future trends in the number of potentially eligible persons depend upon factors such as demographics, economic conditions, and the eligibility and generosity of other income-support programs. Therefore, we will utilize various databases that include this information.

For demographic and economic information, one useful data set is the Current Population Survey (CPS), which is produced monthly by the Bureau of Labor Statistics. This data set samples over 50,000 households and includes information about age, occupation, industry, hours worked, earnings, and sources of income. It also includes responses to the question, "Do you have a health problem or disability which prevents you from working or which limits the kind or the amount of work you can do?" It is the largest and most comprehensive survey collected by the government on a frequent basis, and it is the basis for unemployment, employment, and population statistics for the U.S. economy.

Participation in other income-support programs is an important factor in who may apply for SSDI and SSI programs. While the CPS offers information about income from other sources, the Survey of Income and Program Participation (SIPP) offers more detailed information about the interaction of various income-support programs. The SIPP is a longitudinal survey that has followed 20,000 families since 1984. SIPP also contains demographic and employment information, as well as questions regarding impairments that may limit work, and information regarding ADLs and IADLs.

Since these data sets potentially offer long time-series of information (although usually with a tradeoff against number of observations), we can pursue various methodologies to project past trends into the future. These methodologies would include the standard forecasting models used by the Bureau of Labor Statistics (BLS) as they project occupation and industrial trends into the future. The current 1998 *Occupational Outlook Handbook*, prepared by the BLS, includes detailed occupation and industry employment trends through the year 2006. The model projects GDP and the distribution of GDP by its major demand components. Industry output is converted to industry employment, and then through a detailed industry-occupation matrix, projections are made for 261 industries and 510 occupations. We will then overlay on these projections estimates of the age and impairment distribution within these occupations, obtained from the CPS, or other sources. Other factors, such as workplace accommodations and safety initiatives, and the characteristics of other income-support programs, will also be incorporated into the model by assumption, as informed by data developed in the DES.

MICROSIMULATION OF SSDI AND SSI MODEL RESULTS

Our objective in testing complex option value and dynamic programming models is to estimate parameters that explain observable behavior in a structural manner. The application models specified above will allow us to simulate three particular types of policy changes using these models; increases or decreases in the probability of acceptance onto SSDI and SSI rolls, increases or decreases in the level of benefits, and greater accommodation of workers with disabilities.

Changing the probability of acceptance can be simulated by reducing all probabilities for all states, since we will be using different probabilities for each state and year. Based on existing research, people are believed to make reasonable predictions of the probability of acceptance by a social welfare program. If it is perceived that the eligibility requirements are being interpreted more or less strictly, this will lead to changes in the number of applicants (Parsons, 1991b).

Increases or reductions in SSDI and SSI benefits change the relative attractiveness of alternatives to work, holding constant other disability determinants. Between 1989 and 1993 the beneficiary population grew by nearly 30 percent. Our models can be used to estimate how changes in benefits would change both the SSDI and SSI populations and overall disability expenditures. They can also predict the marginal effect such a policy change would have on increasing the employment and labor earnings of persons with disabilities.

Greater accommodation can be thought of as an attempt by employers to augment the capital that people with disabilities use on a job to increase their productivity. This should make them more productive workers and increase their wages, in line with conventional human capital models. Like other forms of capital, investments in accommodation are costly. The ADA, however, requires employers to pay this cost, since workers cannot be charged for the accommodations that are required by law, and their wages must be the same as workers who are not accommodated. An increase in accommodation could be simulated by randomly assigning additional persons accommodated status, or assigning an increase in the tendency for persons of a given health condition, age, sex, education, or marital status to be accommodated. These decisions will be based on the most current literature and advice from our expert consultants and SSA staff.

Microsimulation greatly extends the range of inferences and implications that can be derived from the estimated parameters of a model. The basic process consists of fixing the values of a set of variables found in the model (typically, the exogenous variables) and randomizing over one or more of the remaining variables (typically the endogenous variables) in the model. The exercise can be carried out for a single individual whose characteristics represent an important or interesting case (i.e., a "representative individual") or for a sample of individuals.

While virtually all models permit us to determine the expected values of key outcomes under the hypothesized model structure, microsimulation offers the additional advantage of producing an entire distribution of outcomes upon which the expectation is based. This distribution can arise due to both systematic and stochastic elements of the model. Thus, using microsimulation, we can alter the observed values of exogenous variables and calculate the distribution of behavioral responses to these changes.

In addition, microsimulation is a powerful tool with which to determine the range of uncertainty associated with model predictions. For example, our estimated model may suggest that a given change in disability policy—perhaps brought about by a particular aspect of the redesign—leads to an X-percent decrease in applications. With microsimulation, we can determine whether the uncertainty associated with this conclusion includes the possibility that the true response to the policy change is zero; taking into account (1) sampling variability in the data that describes the population, (2) sampling variability in the estimated parameters of the model, and (3) stochastic variability inherent in the model.

Estimation of our structural models will produce parameters relating to several sources of stochastic variability. The first represents uncertainty about the future streams of income in each potential state (earnings if employed, earnings if denied, disability benefits, and nonwage income), which is "prediction error" from the viewpoint of the potential applicant. This uncertainty corresponds to the error variances in the earnings autoregressions. The second source of stochastic variability is represented by the structural parameters, which represent unmeasured aspects of utility functions (i.e., unobserved differences between otherwise identical individuals). A third source of uncertainty associated with predictions from the models is not intrinsic to the model itself, but to the procedures used to obtain estimates of its parameters. It is represented by the covariance matrix of the estimated parameters (i.e., the sampling errors of the estimates). We will take account of all three sources of stochastic variation in our microsimulation analyses.

Our microsimulation efforts can be described as a sequence of progressively more complex sampling experiments. First, for a representative individual described by fixed values of exogenous characteristics as well as prespecified values of random utility parameters, we will draw a large sample of future streams of income in each state, sampling from the estimated error-variance distributions of the earnings autoregressions. Note that given a particular draw from these distributions, the decision to apply for benefits is **deterministic**, but the random nature of actual future income makes the future behavior of our representative individual **stochastic**. Stochastic uncertainty about the ultimate decision to apply for disability benefits will grow as we consider longer and longer periods into the future. This sampling experiment represents randomness in the decision process as experienced by the individual decisionmaker.

Second, for a representative individual we will fix observable characteristics but randomize over the random utility terms, sampling from the distributions of these parameters. This allows us to describe the entire distribution of responses to a given environment exhibited by persons with specified characteristics, rather than only their mean. This is important, since there can be a distribution of times from onset of disability to the decision (if any) to apply for disability benefits. Since the model is highly nonlinear, the extent and timing of application behavior may differ across individuals according to several covariate values.

Two sources of stochastic variability intrinsic to the model have been discussed. But these two sources of variability may interact. Accordingly, we plan to examine each separately, as well as the two jointly, in order to assess whether uncertainty in predictions from the model is additive across sources of variability, or otherwise. We are unaware of any prior applications of these methods.

The third sampling experiment is for a sample of representative individuals. We can conduct either of the preceding exercises, randomizing over either the autoregression errors or the random-utility parameters. Of particular interest is the latter exercise, since it leads to information about the distribution of responses to a given policy regime in the population; variability which arises due to unmeasured differences across individuals. Bootstrapping a large (e.g., many replications per individual) sample of draws from the ex post distribution of random-utility parameters will permit us to quantify the uncertainty in an inference on the expected response to a given policy environment. We will use the weighted survey data as our sample in these experiments, using microsimulation to project each sample member's application and beneficiary history forward to age 65.

SPECIFIC MICROSIMULATIONS TO BE PERFORMED

Just as the econometric models described earlier represent the ***dynamics*** of the process by which individuals end up receiving disability benefits, our microsimulations will produce outcomes representing a dynamic process. The two principal ingredients in a dynamic microsimulation are (1) the data base representing the starting point, or initial conditions, of the simulation exercise, and (2) the "rules of motion" embodied in the equations that predict dynamic outcomes. Assumptions pertaining to such factors as future macroeconomic growth, or improvements in population health, can be viewed as aspects of the rules of motion of the process.

As indicated above, we will have three major sets of econometric equations, one representing the reduced-form model, and two representing different types of structural models (option value and dynamic programming). We can also identify a set of microsimulation types that differ according to the nature of the initial conditions assumed.

(1) Simulations for illustrative individuals.

The simplest type of simulations to be performed will represent the simulated lifetime (or, remaining lifetime) of one or more "illustrative individuals." This term is deliberately chosen to be distinct from the "representative" (i.e., average) individual often employed in discussions of Social Security policy. Instead, we will specify the initial conditions in terms of an individual, for example a white male, 25 years old, with 12 years of education, employed at a given wage, and so on. We will obviously have to fix values for every variable that appears in the predictive models. We will then simulate the remaining lifetime (**up to age 65**) of work, disability benefit application behavior and outcomes, and programmatic benefits, for the stated individual. A large number of independent replications of the simulation for a given set of initial conditions will be performed, in order to represent the full range of stochastic variability and hence uncertainty associated with the future of such a person.

The exercise will be repeated for a number of illustrative individuals, chosen so as to illustrate the range of situations of interest. The value of such a simulation is that it can readily be related to a specific set of circumstances, and shows the extent to which disability profiles reflect both deterministic and random elements. It also illustrates the tremendous range of uncertainty concerning whether, and when, a disabling condition will arise over the lifetime, and the uncertain consequences of such conditions. Furthermore, it is possible to use the "illustrative individual" approach to isolate the partial effects of selected variables, holding constant other variables. For example, we can show how education, wage level, or job tenure, among other factors, affect the chances of receiving SSDI or SSI benefits in the future, holding other initial conditions constant. Finally, the initial age can be varied, showing how the chances of receiving disability benefits in the future change for different ages at onset of disability.

(2) Cohort simulations.

The next level of complexity generalizes from a set of "illustrative individuals" to the cohorts from which they come. Here, we will create a data base representing the full distribution of relevant characteristics (i.e., all variables found in the predictive models) in a hypothetical cohort, for example persons age 25 in the year 2000. Whereas the output from the "illustrative individual" described above can only be viewed as illustrative, the output from a cohort simulation is intended to be representative.

In order to do a cohort simulation it will be necessary to assemble a sufficiently large database with which to obtain statistically precise summary statistics representing simulated outcomes. The DES data will play an important role in creating this data base, but is not expected to be large enough to represent, say, those age 20-24 (or 25-29) in the baseline year. We will create such a file by pooling data from independent sources (such as nonoverlapping rotations from a series of CPS samples, for example), using matching or other data imputation techniques as necessary to create the full set of initial values needed to start the simulation.

(3) Residual lifetime simulations.

The two preceding types of microsimulations go part of the way to addressing Research Question 2, namely determining how many individuals not presently receiving disability benefits may apply in the future. A fuller answer can be obtained by simulating the remaining lifetimes **(i.e., to age 65)** of a representative cross-section of the population. For this project the most obvious data file to use as the "starting population" for such a simulation is the DES sample, since it will contain all the information needed to estimate the econometric models and therefore to make predictions from the models. The difference between the cohort simulation discussed above and the "residual lifetime" simulations is that the latter represents individuals from many different cohorts, initially captured at different points in their life cycles.

Thus, we can conduct a series of microsimulations for a sample of individuals, and under a range of hypothetical policy alternatives. The range of possibilities is quite large. However, we will be guided by the policy requirements of SSA and the specifications of the models we estimate. We will focus on changes in accommodation, the level of benefits, and the probability of acceptance, given application for disability benefits. Each type of hypothetical change corresponds to a change in the value of one or more decision criteria as embodied in the model structure. We anticipate that response functions will be highly nonmonotonic in view of the many nonlinearities in the model and the bounded nature of responses. We can, however, extend the range of outcomes studied by randomly assigning applicants to beneficiary status in accordance with the probability-of-acceptance parameters used in the calculations. Thus, we can compute, for an individual or a sample of individuals, the chances of a claim being allowed or denied.

Since a future stream of disability benefits is calculated in the process of evaluating the decision criteria, we will also be able to calculate program benefits. Thus, we can estimate lifetime (i.e., up to age 65) disability benefit streams in each microsimulation experiment as well. Note that in experiments that entail an alternative value of the probability-of-acceptance parameter, the simulated responses include both the induced behavioral response to the changed parameter and the change in the prevalence of beneficiary status given the decision to apply. Thus, a wide range of policy-relevant outcomes can be generated by our microsimulation experiments.

Finally, we note that each of the four successive "levels" of microsimulation analysis described above can be carried out under an additional source of stochastic variation, namely that associated with sampling error in estimates of model parameters. Appealing to the asymptotic properties of the likelihood estimators, we will take the estimated covariance matrix of parameters as defining a population from which parameter vectors can be sampled. Single-factor and multifactorial experiments, in which the selected sources of variability are taken into account, will be carried out in order to derive interval as well as point estimates of the expected effects of disability on SSDI and SSI application behavior and its consequences.

C. Objective 3 - Identifying Factors that Enable Disabled Persons to Remain in the Workforce

Why do some persons with severe impairment continue to work, while others with apparently the same level of impairment leave work and apply for disability benefits? This question is of particular interest to the SSA and policymakers, and forms another of the four core objectives of the research agenda for the DES. Several answers have already been offered (e.g., education level, economic conditions, and social networks. (See Rupp and Stapleton, 1998.) However, much more promising is the possibility that such factors as accommodation at work or home, vocational rehabilitation interventions, specific medications, or innovative employment strategies may assist some persons with disabilities to remain in the workforce, or return to work. If we can discover in the DES the particular accommodations, interventions, and treatments which permit persons with impairments to remain in the workforce, we may be able to recommend strategies and policy initiatives that will help to keep people with disabilities at work. Such an outcome would not only have beneficial cost implications, but it would also improve the quality of life for many persons with disabilities.

This section describes the process the Westat team will use to investigate this important question. Beginning with a review of the literature, we identify those factors already shown to be predictive of persons with disabilities who are working. We then go on to describe our proposed analysis of the DES to elucidate this issue further. We present two approaches: the first is a descriptive approach; the second is a structural modeling approach.

BACKGROUND LITERATURE AND ANALYSIS

Evidence from the New Beneficiary Follow-up Survey (NBF) reveals that approximately 12 percent of SSDI beneficiaries start a job while they are entitled to disability benefits. Another 9 percent medically recover from their disability. However, at least 40 percent remain nonworking beneficiaries until they reach 65 and their SSDI benefits are converted to retirement benefits. The mean time from entitlement to first job is 3.4 years; from entitlement to retirement termination is 6.3 years; and from entitlement to recovery termination is 2.4 years.

The SSDI program contains specific provisions to assist beneficiaries in meeting the conditions of SGA termination. The Vocational Rehabilitation (VR) Reimbursement Program provides rehabilitation services including physical therapy, vocational training, job counseling, general education, and job placement assistance. The SSDI program also has work-incentive (WI) provisions. The trial work period (TWP) allows the beneficiary to work for 9 months while maintaining eligibility for benefits. If the beneficiary successfully completes a TWP (works for 9 months and earns more than $75/month after deducting disability-related [IRWE] costs), the beneficiary then enters the next phase–extended period of eligibility (EPE). This is a 36-month period which provides for automatic reinstatement of their monthly benefit in any month where their work is not above the SGA level minus IRWE costs. (Hennessey, 1996)

Hennessey (1997) investigated the factors that affect the work efforts of disabled-worker beneficiaries. The variables that he explores include: gender, age, race, education, primary insurance amount, vocational rehabilitation efforts, work-incentive provisions, job accommodations, motivation of working, and type of job. He relates these factors to beneficiaries' tendencies to start work and continue to work. Based on the NBF survey, he finds that beneficiaries with vocational rehabilitation services have a higher tendency to start working and a lower tendency to stop; those with vocational training or general education also have a higher tendency to start work, but this does not influence their tendency to continue working. Beneficiaries with job placement assistance are more likely to start work but also more likely to stop work.

Some work has already been conducted to determine why some disabled individuals work and others do not. In 1996, the Government Accounting Office (GAO) conducted a small study of 69 persons who were receiving SSDI benefits and working in the Washington, DC, Atlanta, and San Francisco metropolitan areas (GAO/HEHS, 1998). The typical beneficiary reported working 28 hours each week and receiving about $10.60 an hour. Most reported that financial need and the desire to enhance self-esteem were the main reasons for attempting work. They indicated that a range of factors enabled them to return to work including improved functioning and encouragement from family, friends, health care providers, and coworkers. Contributors included high self-motivation and a flexible work schedule that allowed them to receive health care services, job-related training and vocational rehabilitation services.

About four in ten respondents said they planned to leave the SSDI rolls in the future. Availability of worksite-based health insurance appeared to delineate respondents who planned to leave the rolls in the future from respondents who planned to stay. Other factors mentioned included religious faith, job coaches, assistive devices and equipment, and enablements from the Americans with Disabilities Act. Many respondents reported that they had experienced impediments to employment such as limited skills and training, or employers' failure to recognize their abilities.

Yelin and Katz (1994a) analyzed the National Health Interview Survey for the years 1970 to 1992 to estimate labor force participation rates among persons with and without disabilities in each year and to compare the change in the rates for the two groups over time. In describing labor force participation rates, they compiled separate analyses for men and women and partitioned the data into three age groups—18-44 years ("young" workers), 45-54 years (those in the "prime" of a career), and 55-64 years ("older" workers). Employment trends from 1970 to 1992 indicated that the labor force participation of persons with disabilities was tied to overall labor market dynamics, in both the long and the short term. In the long term, a decline in the labor force participation of men over the period—particularly older men—was more pronounced among men with disabilities. By contrast, an increase in the labor force participation of women during the same period—especially younger women—benefited women with disabilities. In the short term, persons with disabilities experienced proportionally larger gains during periods of labor market expansion than did those without disabilities, but suffered proportionally greater losses during times of contraction than did their non-disabled counterparts.

Using data from the March Supplement to the Current Population Survey for 1981 through 1993, Yelin and Katz (1994b) concluded that contractions in the labor market constrain opportunities among persons with disabilities, setting boundaries on their employment prospects. They also concluded that individuals' specific work histories largely determine whether they will retain, lose, or find jobs. In 1993, the probability that the typical person with disabilities was working the week prior to interview was 34 percent, or about half the rate among persons without disabilities. However, for those persons with disabilities and a work history, the gap is narrowed. About half of such persons worked in the week prior to interview, or more than 16 times the rate for persons with disabilities and no prior work history. Overall, about a quarter of all those working in the prior year were not working in the week prior to interview. Among persons with disabilities, 56 percent of those who had worked in the prior year were no longer doing so, and they were 2.3 times more likely to have stopped working than persons without disabilities.

The employment situation of the individual can also alter the probability of job loss. Persons with disabilities in the best combination of occupation and industry were only about 37 percent as likely to stop working as those in the worst, 98 percent of whom were not working in the week prior to interview. Union members with disabilities working the year prior to interview were less than half as likely as nonunion members with disabilities to stop working by the week prior to interview (27 versus 57 percent).

Westat conducted a preliminary analyses of the issue of those who are disabled but continue to work. We asked the question, "What are the factors that predict why some persons with disabilities work while others do not?" Besides demographic variables shown in the literature to be related to working (e.g., age, education level), we hypothesized that social factors (e.g., living alone) and accommodations would also be strong predictors (e.g., use of mobility aids).

Using data from the 1994 National Health Interview Survey Disability Supplement (Phase One, 1994) on persons age 18 to 65, we used multiple logistic regression to examine predictors of work for those with mental and physical disability. We chose as the dependent variable the inability to work and also not currently working vs. the other possibilities (able to work/working; able to work/not working; unable to work/ working). Independent variables were age group, sex, race, geographic region, living alone, and mobility aids (for the physical function model). Among those with a mental disability, persons with a grade 8 education or less were 17.6 times likelier than those with the highest education level to be unable to work and not working. Those living alone were 1.9 times likelier to be unable to work and not working compared to those not living alone. When assistive aids were added to the model for disability in physical functioning, the effect was a 2.1 fold increase in the likelihood of being unable to work/not working.

This analysis identified the importance of education level and social factors in assisting persons with disabilities in staying at work. For the group of physically disabled persons identified, mobility aids appeared to be acting as an indicator of severity. However, with further analyses that contained other measures of accommodation, (such as those which will be collected in the DES) and controlling for severity of disability, we hypothesize that accommodation and intervention will be more predictive of working among persons with disabilities.

A modest literature has developed exploring the effectiveness of financial incentives in encouraging work among the disabled (or conversely, how financial disincentives discourage work). The results of a sample of these studies are reported in Table 11**.** In general, the studies report that work effort among the disabled is not particularly responsive to financial incentives. After reviewing the literature on financial incentives that promote work among welfare recipients and its possible application to understanding the response of the disabled to similar incentives, Hoynes and Moffit argue that the potential effect of financial incentives on the work effort of the disabled is small.

Furthermore, training programs aimed at returning disabled individuals back to work have met with limited success. For example, an SSA demonstration project, Project Network, assigned disabled individuals to treatment and control groups to assess the effectiveness of a particular return to work training program. Less than 3 percent of those enrolled in the program were successfully placed in jobs.

Furthermore, successful job placements may be subsequently undermined by the poor health of the disabled individual. An evaluation of a return to work program based in Ontario, Canada, reports that 40 percent of those initially placed in a job had to leave within 2 years due to health problems related to their disability (Haveman and Wolfe, 1999). All of the above suggest that a simple one dimensional approach to encouraging work among the disabled is unlikely to succeed because of the manifold and varied barriers to employment.

Table 11. Summary of Labor Force Effects of Disability-Related Benefit Levels

Study	Data Set	Sample Analyzed	Results
Leonard (1979)	Social Security Survey of Health and Working Conditions	1,685 men aged 45–54 in 1972	Elasticity of beneficiary status = .35
Parsons (1980a)	National Longitudinal Survey	3,219 men aged 48–62 in 1969	Elasticity of labor force nonparticipation = .63
Parsons (1980b)	Panel Study of Income Dynamics		Elasticity of labor force nonparticipation = 1.80
Slade (1984)	Longitudinal Retirement History Survey	5,403 men aged 58–63 in 1969	Elasticity of labor force nonparticipation = .81
Haveman and Wolfe (1984b)	Panel Study of Income Dynamics	741 men aged 45–62 in 1978	Elasticity of labor force nonparticipation = .06 to .21
Haveman, de Jong, and Wolfe (1991)	Panel Study of Income Dynamics		Elasticity of labor force nonparticipation = .97 (single female heads) = .23 (married women)
Gruber (1996)	Canadian Labor Force Survey	Men 45-59 1985-89	Elasticity of labor force nonparticipation = .25 to .32

Source: Haveman and Wolfe, 1999.

It is widely recognized that a disability may directly interfere with an individual's ability to perform work-related tasks. For example, the recent passage of the American with Disabilities Act reflects, in part, a belief that increased workplace accommodations would create an environment in which a greater proportion of disabled individuals might return to work. But it is less often recognized that a disability may also create secondary barriers to return to work. Examples of such "secondary" barriers produced by disability are increased transportation costs, increased job search costs, and decreased access to social networks related to the provision of job leads.

In addition to "secondary" barriers produced by a disability, disabled individuals may also face barriers to employment that result from low socio-economic status. The employment effect of barriers related to low socio-economic status is studied in the welfare-to-work literature, which is briefly reviewed here. Another reason for considering the welfare-to-work literature is that these same barriers to work are, in many cases, analogous to the "secondary" barriers produced by a disabling event. Thus, the results of this literature are suggestive of the effects that low socio-economic status and "secondary" barriers may have on employment outcomes for disabled individuals.

A number of studies have pointed out that many welfare recipients are unable to retain entry level jobs, causing them to return to public assistance (Harris, 1996, Pavetti, 1993, Blank, 1997). This "cycling" between public assistance and employment results, in part, from the high qualifications demanded by employers for entry level jobs and by the inability of many welfare recipients to meet those requirements (Holzer, 1996). A tight labor market may induce an employer to hire a welfare recipient who does not have the necessary qualifications, but they may replace that person when another more qualified worker becomes available or when labor market conditions ease.

A number of studies have concluded that many welfare recipients have little or no experience performing basic work tasks that are routinely required of entry-level positions (Band and Ellwood, 1994; Harris, 1996). Furthermore, some welfare recipients lack the social skills necessary to work productively with co-workers and superiors (Blank, 1997). Another barrier to successful integration into the labor market is discrimination. Turner (1991) documented through a series of audit studies that discrimination against racial minorities is common in the application process for entry-level positions. The same is believed true for persons with disabilities.

There are a number of health-related barriers that negatively affect the chances of welfare recipients in the labor market. Welfare recipients are more likely to be victims of physical abuse than the general population, are more likely to suffer from mental disorders, especially depression, and more likely to abuse drugs or alcohol (Olson and Pavetti, 1996). Furthermore, the poor health that plagues many welfare recipients and that may lead to disability has been shown in a number of studies to directly and negatively affect their employment prospects (Olson and Pavetti, 1996; Bird and Freemont, 1991).

The spatial concentration of poor people, including welfare recipients, into high poverty areas that are distant from job opportunities also limits their ability to find job vacancies. There is a large literature on the employment effect of spatial mismatch, which is reviewed by Kain (1992) and Ihlanfeldt and Sjoquist (1999). Several empirical studies have reported quantitatively large and statistically significant associations between spatial mismatch and negative employment outcomes for urban minorities. As the costs of searching for jobs increase, either by increasing the physical distance between the unemployed and vacancies or through discriminatory practices, the job prospects for welfare recipients deteriorate.

Finally, Danziger, et al. (1999) analyze the cumulative effect of these types of employment barriers using a sample of 3200 welfare recipients from Michigan. The primary findings from this study are reported in Tables 12 and 13. The first column of Table 12 indicates that as the number of barriers to employment mount, the drop in the probability of employment accelerates. Thus, those with multiple barriers have much less chance of obtaining a job. This is clearly relevant for persons with disabilities, who often face multiple barriers to employment.

The individual impact of employment barriers is reported in Table 13. The estimates indicate that, as expected, educational attainments and employment skills are important factors in the success of welfare recipients. However, the results also indicate that the quantitative impact of transportation problems, health problems, and discrimination is nearly the same as education and work experience. Again, these findings are highly relevant for persons with disabilities.

Table 12. Cumulative Effect of Employment Barriers on the Probability of Employment for Former Welfare Recipients.

Number of Barriers	Probability of Working 20+ Hours Per Week
0	80.5
1	71.3
2 – 3	61.7
4 – 6	40.8
7 or more	5.7

Source: Danziger, et al. (1999)

Table 13. Quantitative Effect of Individual Employment Barriers on the Probability of Employment for Former Welfare Recipients

Barriers	Probability of Working 20+ Hours Per Week
None	81.6
Less than HS Education	70.9
Fewer than 4 job skills	64.4
Perceived Discrimination	68.5
Transportation Problem	67.4
Major Depressive Disorder	73.1
Drug Dependence	60.2
Mother's Health Problem	68.7

Source: Danziger, et al. (1999)

WELFARE TO WORK POLICY CHANGES

The last several years have witnessed dramatic policy changes in the welfare system. Significant reform was first initiated in a number of states during the early 1990s. Wisconsin was one of the early states to switch from an entitlement-based welfare system to one in which welfare recipients were expected to assume greater responsibility for finding work and gaining economic self-sufficiency.

In 1996, Congress galvanized many of the reform principles initiated by Wisconsin and other states by passing the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA). The Act replaced the old welfare system (AFDC) with Temporary Assistance for Needy Families (TANF), which focused on work and responsibility and provided States with flexibility to create the best approaches for their individual circumstances. The new law contains strong work requirements, places a time limit on most assistance, reduces welfare dependency, and encourages two-parent families.

Since this new strategy has been implemented, there has been a dramatic decline in the welfare caseloads. The number of recipients fell from 14.1 million in January 1993 to 7.3 million in March 1999, a decline of nearly 48 percent. An estimated 2.7 percent of the population received assistance in March 1999, the lowest percentage since 1968. Three-quarters of that decline took place after the federal welfare law went into effect in 1996 (U.S.D.H.H.S, 1999, Second Annual Report to Congress).

Several studies have shown that in recent years the implementation of welfare reform is the single most important factor contributing to the widespread and continuous decline in caseloads during this period. (Blank, 1997; Ziliak, et al., 1997; Executive Office of the President of the United States, Council of Economic Advisers, 1999.) Nonetheless, the strong economy since 1993 has also played a large role in the reduction in caseloads. Between 1993 and 1996, 26 to 36 percent of the decline was due to the improved labor market, whereas less than 15 percent was due to changes in the welfare system. After 1996, these two effects switch in importance with welfare reform accounting for 33 percent of the caseload reduction, whereas the economy accounted for only 10 percent of the decline.

Understanding the reasons for the reduction in caseloads is hampered by state reporting methods. State reporting forms list six categories of reasons for closing caseloads: employment, marriage, five-year limit, sanctions, state policy, and other. Unfortunately, many of the case closures are classified as “other.” This category includes a variety of unknown reasons, such as the family volunteering to close the case. According to the most recent state reports, 56 percent of the cases were closed due to “other” reasons, 21 percent due to employment, 16 percent due to state policy, and 6 percent due to sanctions. Some analysts believe that these reports actually underestimate the number of welfare recipients leaving the program because of employment.[8]

[8] Source: National Emergency TANF Datafile as of 12/9/98, Table 30,

More detailed state- and county-level analyses suggest that between 50 and 65 percent of former TANF recipients found work immediately after leaving TANF. However, some of these former welfare recipients returned to the welfare rolls within a short period of time. As a result, the proportion of recipients who are working (including employment, work experience, and community service) in a typical month is much smaller, about 25 percent in 1997. Nonetheless, this proportion of recipients who are working represents nearly a fourfold increase over the 7 percent recorded in 1992 (U.S.D.H.H.S, 1999, Appendix 14:1).

Results from a survey of employers in Michigan reveal that absenteeism, attitude problems, basic skills, job skills, substance abuse, and crime, (in that order of importance) were the problems employers encountered with welfare recipients they hired (Holzer, 1999). The same survey revealed that employers are willing to provide some assistance to help welfare recipients overcome some of the barriers. Eighty percent of the 900 Michigan establishments that responded to the survey said that they would help provide job skills to welfare recipients. However, only 33 percent said that they would provide basic skill remediation, which is one of the first hurdles facing harder-to-serve welfare recipients. Further, half of the employers expect either tax credits or technical assistance from the government in order to provide training. Employers were much less willing to help with child care and transportation. Less than 10 percent responded that they would definitely help provide child care and transportation and about the same percentage responded that maybe they would help provide these two services.

To a certain extent, state-provided services complement employer-provided services. State work-first type programs focus primarily on providing child care and transportation. Some training is available, but only after the person is already working. Most of the effort is directed toward job search assistance. However, States can and do also use their federal and state funds for non-traditional welfare services, such as non-medical substance abuse and domestic violence services.

With the significant reduction in welfare caseloads, the remaining welfare cases typically include individuals who have significant barriers to overcome and consequently are harder to serve. Furthermore, an unknown number of closed cases includes those who are not able to meet the work requirement or even the job search requirement because of physical or mental problems (i.e., the disabled). These individuals, who were once supported by the welfare system, may now seek assistance from other sources including social security disability programs.

One of the major contributions of the DES will be the documentation of similar types of barriers among the disabled. Questions regarding the employment histories can be used to access an individual's employment skills. Questions regarding mode of transportation used for work will be used to assess the extent to which individuals experience transportation problems. When these and other questions assessing the prevalence of social barriers are coupled with the extensive medical histories n the DES, researchers will be able to assess in detail the ways in which health conditions and social constraints interact to prevent disabled individuals from achieving successful employment.

MODELING APPROACHES

Since return-to-work or stay-at-work options are believed to have dramatic impacts on program expenditures, there has been both public and private sector interest in this subject (GAO, 1996; GAO, 1997; Hunt, et. al., 1996). However, very little is known about the actual potential of work among SSDI or SSI beneficiaries. The large sample of persons with severe disabilities in the DES, the distribution of work patterns that will be observed in the sample, and the relatively full description of accommodations, rehabilitation interventions, and employment strategies promise the best opportunity yet to determine the actual work potential of persons with significant disabilities.

We will begin with a descriptive approach using statistical methodologies to answer the research question: "What are the factors that predict why some persons with disabilities work while others do not?" The first technique is multiple logistic regression analysis. The dependent variable will be a binary variable (working/not working). The independent variables (see table 7) will include the following categories: demographic factors (age, sex, geographic region, occupation), personal economic factors (e.g., current wages/income or wages/income at the time the individual stopped working, family resources), accommodation factors (e.g., engineering, administrative, independent living), medical interventions, and rehabilitative efforts (e.g., vocational training, occupational, or physical therapy). The second technique employed will be CART to obtain a branching classification tree which identifies associations and interactions. We are also exploring other analytical methods such as neural networks that could offer new insights into the complex relationships between disability and work. Our descriptive approach to the research questions will utilize whichever of these techniques is more informative and understandable.

In addition, we will construct structural models similar to those discussed earlier (in Section II-B, where they were used for determining future trends in SSA applicants). The underlying premise is that at least some individuals choose when to leave the labor force and apply for disability benefits, given their underlying impairment. The same approach can also be applied to the decision to leave work or stay and seek accommodation. To our knowledge, such models have not been used in this way previously.

By studying the current work behaviors of persons who would be eligible for disability benefits (according to the DES simulated disability determination process) if they applied, we will gain valuable insight into the determinants of work effort and labor market success. The logistic regression analysis will identify differences between two groups of persons with severe disabilities, those who work and those who do not at the time of the survey, controlling for health condition and functional ability. This should allow sorting of influences such as wage level, replacement level of income support benefit, rehabilitation experience, employer accommodations, treatments, and other factors. Since the DES will have the best data yet available, we expect to find significant differences even if the sample size is not great.

By comparing the work behaviors of those who are moderately disabled with those who are severely disabled (according to the DES medical examination and functional analysis), we will be able to achieve a better understanding of the potential work contributions of the disabled, assuming their disabilities could be more effectively accommodated. In other words, if employer accommodation of specific disabilities could be improved, or more effective treatment options could be developed for specific disabling conditions, what kind of employment results could be predicted? One way to answer this question is by comparing individuals who are slightly different in their impairment level. This approach can be extended to the entire range of disabilities observed in the DES sample. It would be reasonable to hypothesize that marginal improvement in treatment or accommodation could produce marginal improvements in employability. Constructing comparison groups internal to the DES will allow the development of such measures.

By reviewing SSA administrative records matched to those participants in the DES who grant access to such records, it should be possible to develop a small sub-sample of individuals from the DES who have attempted to return to work under various SSA programmatic provisions. These individuals could be matched to others with similar disabilities to develop a picture of the determinants of successful return-to-work efforts. For instance, a sub-sample of beneficiaries who have used the trial work period might be accumulated for analysis. By examining the experience of such a sub-sample in returning to work, successful or otherwise, additional insight into the barriers to re-employment for beneficiaries could be gained. It is not clear how many such individuals will be available for analysis in the DES, but with the cooperation of SSA and sample participants, we will analyze the data that are available.

The coincidence of the increase in the SGA level from $500 to $700 effective 1 July 1999 creates another opportunity to study the response of DES sampled individuals to financial incentives. Since the change occurred at a discrete point in time and applied to all SSDI beneficiaries, we can use this "natural experiment" to compare earnings during the first half of 1999 ($500 SGA limit) with earnings during the second half of 1999 ($700 SGA limit). This will provide a rather direct test of the responsiveness of labor supply among SSA beneficiaries. If the SGA limit is serving to artificially restrict hours of work among a significant number of claimants, this should be apparent from the indicated comparison.

In short, the DES will enable the most detailed examination to date of impairments and their resulting labor market outcomes. Our analyses are intended to exploit the DES data to develop an improved understanding of the potential of accommodation and other interventions to improve the employment prospects for persons with disabilities. It would be logical to expect that additional policy options and insights might emerge from these analyses. Understanding the "why" of a problem is usually the first step in designing the "how" of addressing that problem. Our analysis of the DES will seek opportunities to promote employment and overcome barriers to employment for persons with disabilities.

D. Objective 4 - Identifying Self-Reported Measures Appropriate to Estimate Program Eligibility

In most social science surveys of disability, the definition of disability is operationalized using survey questions that ask whether the individual has some type of work limitation (e.g., “Do you have any impairment or health condition that limits the amount or kind of paid work you can do?” and “Does any impairment or health problem limit the kind or amount of work you can do around the house?”). The problems inherent in this type of question are well-documented (see Parsons, 1980; Bound, 1991). Still, these measures are highly correlated with more objective measures of health status (see Bound, 1991; Stern, 1989). If self-reported data can be shown to be useful in accurately predicting SSDI and SSI status, perhaps one of the standard longitudinal survey efforts already existing could be adapted to meet the need of providing self-reported measures sufficient to estimate future program eligibility.

One of the exciting features of the DES design is the opportunity to directly test the accuracy of self-reported data in determining disability status and program eligibility. Since respondents will be asked to self-report their disability status in the questionnaires, these responses will be available for a random sample of the population. Then, it will be possible to analyze how, and with what eligibility consequences, the self-reported responses distort true disability status as discovered in the subsequent examination of the respondent, including the physical examination, functional assessment, medical evidence of record, and other evidence.

It is known that self-reported work limitations are highly correlated with reports of either chronic health conditions or functional limitations. Burkhauser and Daly (1997), using the 1986 Health Supplement to the Panel Study of Income Dynamics (PSID), showed that men and women who responded “yes” in two consecutive waves of data for health-related work limitations were in poorer health and were more likely to have functional limitations than either individuals without such limitations or individuals who reported “yes” to limitations in only one wave. They also found that those with longer term health-related work limitations were less likely to work and had lower median labor earnings and household income than the others. Burkhauser and Wittenburg (1996) found the same results using the health supplement of the 1990 Survey of Income and Program Participation (SIPP).

While the DES survey will also employ self-reported measures of health limitations in the screeners, it will contain far more detailed information on severity of disability, functional capacity, and need for accommodation. In addition, some 5,500 DES respondents will have medical exams administered as part of their participation in the study, supplemented by the gathering of medical evidence of record with which to develop the simulated disability application folders. Because the measures of health status in the DES will be far superior to those in other social science surveys (usually only self-reported data), we will be able to identify more accurately those with health conditions which make them eligible, or nearly eligible, to receive SSDI benefits; as well as those with a moderate level of impairment but of a severity not close to making them eligible for SSDI benefits. Then, we will be able to compare predicted disability from the self-report data from the screener and other parts of the survey with the SSA standard for disability determination to seek new approaches in predicting future disability applications from self-reported data and covariates.

If it can be shown that self-reported data provide an adequate estimate of program eligibility, then much more confidence can be put in such data (and in existing sources such as CPS, NHIS, SIPP, PSID, and HRS which collect these data), for analyzing and forecasting the population of future program applicants. If such self-report data do not prove to be accurate, it calls into question the adequacy of self-reported measures of disability that have been used for policy and analytical purposes for many years.

This section discusses how the Westat team will identify the self-reported measures appropriate to estimate disability. Our approach consists of the following steps:

- Develop estimation models for predicting eligibility from self-reported data;
- Determine the sensitivity and specificity of predictive models;
- Identify ongoing surveys which provide similar self-reported data to enable tracking of future trends in the eligible population.

DEVELOPING ESTIMATION MODELS FROM SELF-REPORT DATA

To begin, we will model a variety of self-reported and observational health measures on their own and in combination with one another (Table 14). For example, based on data from the Initial and Follow-up Screeners and Comprehensive Survey Interview, we will develop models which use self-reports of general questions (e.g., “Do you have any impairment or health condition that limits the amount or kind of paid work you can do?” or “Would you say that your health in general is excellent, very good, good, fair, or poor?”), questions on specific health conditions (e.g., those mentioned in the Listings), work limitations, functional abilities, and medical history. Models could also contain data based on interviewer and physician/nurse practitioner/psychiatric social worker observation of the SP’s functional abilities, as well as medical test results, and MER. We will develop models which combine some or all of these types of data (e.g., Models 9 through 15 in Table 14). The results obtained from each model can be compared against the simulated SSA disability determination.

Given that our goal is to predict objective health, we also want to reduce the variance of the disturbance. We propose to include such exogenous variables as age, race, sex, region or state, industry, and occupation. Measures of the local macroeconomic performance, unemployment rates, and other factors affecting job prospects can also affect disability, both directly and through psychological means (e.g., good job prospects can make people feel better, or at least behave as though they do). Such variables will be included through geocoding the respondent's address This will enable matching to a host of social indicators that might also be tested for their influence on self-reported disability.

The resulting equations have no structural or causal interpretation but may be very useful in extrapolating trends in eligibility for disability over time. Predicting disability status ultimately can be done as a probability assigned to each person based on their characteristics, or as a discriminant function, say with probabilities above 0.75 assigned to be a predicted "1" or disabled, and probabilities below 0.75 assigned to be a predicted "0" or healthy. Other arbitrary breakpoints will also be considered, upon discussion with SSA.

Any of the analyses proposed here can be carried out with or without sampling weights, and either is justifiable as a prediction problem. To discuss predictions in the population, however, weights should be applied since the sampling is based on health or disability status, which is the dependent variable. In fitting logistic regression models, a mixture of weighted and unweighted analyses will be used. In evaluating significance of regression coefficients, Westat will use WesVar.

The logistic regression model provides a tool that can be used for predicting disability. The logistic model is given by

$$E(y) = \frac{e^{x'\beta}}{1+e^{x'\beta}},$$

where

y_i = the disability outcome

β = vector of logistic regression coefficients

x_i = vector of self-reported values for the *i*-th individual.

When the regression coefficients have been estimated, a simple rule for assigning disability status to the *i*-th individual would be to predict disability if and only if

$$\frac{e^{x'\hat{\beta}}}{1+e^{x'\hat{\beta}}} \geq 0.75,$$

where $\hat{\beta}$ is the vector of estimated logistic regression coefficients. The assumption we make here is that that if a person has at least a 75 % chance of being disabled, then the person would be eligible for benefits. Westat will work with SSA to identify the most appropriate decision rule to predict benefit entitlement.

The prediction equation can be a logit, but we will also use other statistical modeling techniques which have been shown to do a good job of maximizing the explained sum of squares. Thus, we will consider using semiparametric techniques (these avoid restrictions of functional form and can attain better results than logits). We will also explore Average Derivative Estimation, a technique which is feasible, noniterative, and well established in the econometric literature.

Sensitivity and Specificity

Sensitivity and specificity are the primary measures of how well a model discriminates between two conditions (in this case disabled versus non-disabled) (Kramer, 1988). We will use the simulated SSA disability determination to compare how well our models and cut points predict disability status. If a given model is *sensitive*, then most persons who are truly disabled will be so identified by the model. However, the model may also incorrectly identify disability status for people who are *not* disabled. A model which is *specific* is usually correct when it identifies who is disabled. A model which is sensitive *and* specific predicts disability status for those people who are disabled (sensitivity) and *only* those who are disabled (specificity).

To evaluate the best cut-off for sensitivity and specificity, it will be useful to develop receiver operating characteristics (ROC) curves (Campbell, 1994). These curves plot sensitivity against specificity and can be used to develop optimal cut-offs for classification of eligibility status. For example, it may be determined that

$$\frac{e^{x'\hat{\beta}}}{1+e^{x'\hat{\beta}}} \geq 0.65$$

is a better cut-off when both sensitivity and specificity are considered.

Table 14. Examples of Alternative Models for Predicting Eligibility from DES Components

DES Components	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12	Model l3	Model 14	Model 15
General Questions*	**X**								**X**	**X**	**X**	**X**	**X**	**X**	
Self-reported Health Conditions		**X**							**X**	**X**	**X**	**X**	**X**		**X**
Self-reported Work Limitation			**X**						**X**	**X**	**X**		**X**		
Self-reported Functional Abilities				**X**					**X**						
Results of Performance Tests					**X**				**X**	**X**					**X**
Self-reported Medical History						**X**			**X**					**X**	
Results of Medical Examination							**X**		**X**				**X**		
MER								**X**	**X**				**X**		

*Examples: “Do you have any impairment or health condition that limits the amount or kind of paid work you can do?”; “Would you say that your health in general is excellent, very good, good, fair, or poor?

While sensitivity and specificity give a good indication of how well a model predicts eligibility, there is one disadvantage: equal weight is given to both sensitivity and specificity. In identifying who is eligible for disability benefits, there may be greater "costs" or penalties associated with *not* identifying some who are eligible than there are for incorrectly classifying some persons as eligible – or vice versa. For this reason, it may be useful to apply a *loss function* approach to deciding on the final set of "best" predictor variables (Fishburn, 1988).

The loss function weighs the misclassification errors by the costs of each type of error, allowing an evaluation that reflects possibly unequal costs. Since sensitivity reflects the chance of identifying a truly eligible person incorrectly and specificity reflects the chance of classifying people who are *not* eligible, a simple loss function might be

$$\text{Loss} = c_1 \times (1\text{- sensitivity}) + c_2 \times (1 \text{ - specificity})$$

where c_1 and c_2 reflect the costs of incorrectly misclassifying truly eligible persons and truly non-eligible persons, respectively.

Sensitivity and specificity can be determined for each of the candidate models, using an approximately optimum cut-off. The loss function approach is particularly useful for ranking and selecting the best model. The best prediction model should be the one with the lowest error rates, or the smallest cost associated with misclassification errors.

ONGOING SURVEYS WITH SELF-REPORT DATA

If a group of self-reported measures is shown to have high sensitivity and specificity, a cost-efficient way of periodically predicting program participation on an ongoing basis would be to incorporate those measures in surveys already being conducted. We have identified five separate surveys that gather information on health or disability status and program participation, and that could be considered for this purpose: the Health and Retirement Study (HRS), which was inspired by the 1978 Survey of Disability and Work, the Panel Study of Income Dynamics (PSID), the Current Population Survey (CPS), the Survey of Income and Program Participation (SIPP), and the National Health Interview Survey - Disability Supplement (NHIS-D). This section briefly describes each survey, including the

Survey of Disability and Work, which is no longer being conducted. Members of the Westat team have considerable experience in working with each of these databases.

(1) Survey of Disability and Work

The 1978 Survey of Disability and Work (SDW) is a nationally representative dataset containing information on workers with disabilities. (For technical details, see Bye and Scheckter, 1982.) This survey of the prevalence of work disabilities in the working age population was conducted by the Social Security Administration and contains two sampling frames. The first is a subsample of the National Health Interview Survey (NHIS) and is representative of the general population of noninstitutionalized persons aged 18 to 64. It contains data on 5,652 persons. Because the onset of a health condition in this working age population is a relatively rare event, and applying for disability benefits is even rarer, a second sampling frame was developed to increase the number of respondents who had applied for benefits. This second frame is drawn from Social Security Administration records and consists of 4,207 persons who applied for SSDI.

Respondents in both frames were asked identical questions in 1978, including identifying any health conditions they had, when their main health condition began, and when it first began to limit their ability to work. Additional retrospective information on labor market activity (including occupation, industry, and household characteristics at the time each respondent's health began to limit ability to work) is available. Also available is whether the respondent ever applied for SSDI benefits, and if so, when.

(2) Health and Retirement Study

The Health and Retirement Study (HRS) is a nationally representative sample of men and women born between 1931 and 1941 and their families (Juster and Suzman,1995). In 1992 (Wave 1), a total of 12,654 men and women from 7,607 households were asked detailed questions regarding labor force participation, economic status, and family structure. The HRS contains a disability module (Section J) based on the 1978 SDW retrospective questions which is asked of persons who report a work limiting condition.

While the data set is large, and unlike the NHIS sample is focused on an older working age cohort, even at these ages (51 to 61 in 1992) the presence of disability is still rather low. Thus, the sample size shrinks quickly when questions about specific disability issues are addressed. The HRS sample includes only 750 men between the ages of 20 and 60 at the time their health first began to limit their work, who were employed by someone other than themselves and who were eligible for the SSDI program at the onset of the condition. When the HRS survey data are matched to the Social Security earnings records for respondents who granted the Social Security Administration permission to do so, only about 550 men are available for analysis.

(3) Panel Survey of Income Dynamics

The Panel Survey of Income Dynamics also has some potential for analytical work on disability issues. The PSID is a longitudinal study of a representative sample of 5,000 U.S. households that was begun in 1968. Since the PSID has followed individuals as they have started family units of their own, the number of household units had grown to about 8,700 in 1995. Sponsorship has changed over the years, but currently the National Science Foundation is the major source of funds. General health status and disability status have been core topics from the start, but the major thrust of the PSID is economic and demographic information.

There have been a number of PSID supplements over the years. In the past decade, a series of health supplements for individuals over 55 have been funded by the National Institute on Aging, including general health status, activities of daily living, major illness episodes, and health expenditures. Since the PSID is an annual survey, it would seem to lend itself to measuring emerging disability problems. However, the sample size may be a problem here as well. Burkhauser and Daly (1997) report 366 men and 433 women aged 25 to 61 self-reported a physical or nervous condition that limited the type or amount of work they could do in both 1988 and 1989 in the PSID. They used two years of self-report data to eliminate the effects of short-term illness in their analysis.

(4) Current Population Survey

The Current Population Survey (CPS) is a monthly survey of approximately 50,000 occupied housing units (households) conducted by the Bureau of the Census for the U.S. Department of Labor, Bureau of Labor Statistics. It is designed to gather information about employment and labor force status for a sample representative of the civilian noninstitutional population, aged 16 and over. The monthly unemployment rate for the U.S. economy is derived from this survey. It employs a rotation plan for sample units that provides that three-fourths of sample units are the same from month to month, while one-half are common from the same month a year earlier.

While the CPS is an enormous survey, the interest in disability reflects the possibility that health conditions are limiting labor force participation. It seems unlikely that it could be turned into a tool for accurately measuring the disabled population. Nevertheless, if a very few self-report questions could be developed that would adequately predict future disability benefit applicants, it might be possible to incorporate such questions in the design of the CPS.

(5) Survey of Income and Program Participation

The Survey of Income and Program Participation (SIPP) is a longitudinal household survey focusing on beneficiary status of members of up to 20,000 families, depending upon the year. It began in 1984 and is funded by the Bureau of the Census. SIPP is a panel study, with a rotation structure similar to the Current Population Survey that leads to a series of eight interviews over several years. Particular interest has centered on analysis of the 1990 panel because of the special health and disability focus. Information about general health status is available, plus a selection of ADLs and IADLs, and questions about the utilization of various aids and appliances.

In addition, Lahiri, et al. (1995), were able to analyze a sample with matched SSA information (Form 831) to determine how many of the SIPP participants had applied for disability benefits between 1986 and 1993. This exercise yielded 2,293 applicants from the 1990 SIPP sample. Thus, both the number of observations and the amount of disability specific information available is greater than in either the PSID or HRS.

(6) National Health Interview Survey - Disability Supplement

The National Health Interview Survey (NHIS) - Disability Supplement has two phases. The first phase (administered in 1994 and 1995) was administered at the same time as the NHIS Core and contains questions on physical functioning, sensory impairments, ADLs and IADLs, cognitive and mental symptoms and diagnoses, and diagnosed health conditions. Based on a screener for identifying persons with disabilities, Phase II of the survey (known as the “Disability Followback Survey”) was administered between 1994 and 1997 on those individuals identified as disabled by the screener. The followback survey contains detailed questions on accommodations, treatments, and rehabilitation.

It is possible to combine the core NHIS survey and the Disability Supplement to obtain a weighted estimate of the U.S. population who is disabled. Westat produced several weighted estimates using various definitions of disability. One definition (which included responses to questions on physical functioning, sensory impairments, ADLs, cognitive and mental symptoms and diagnoses, and a selected series of health conditions) found that 9.5% of the population age 18 - 69 is severely disabled and an additional 7.4% may be considered disabled but not severely so. Sample size for this group was 68,123.

Table 15 indicates that the general longitudinal databases described above (excluding HRS and the NHIS - Disability Supplement, which did not yet exist at the 1989-90 point of comparison) yield roughly comparable figures for the prevalence of disability, when they are adjusted to similar populations. It is reassuring, at least, to know that people tend to respond in similar ways to similar questions, even though across different surveys with different designs.

Completion of the DES will not only bring together state-of-the-art social science with a "gold standard" of disability determination, it will also make it possible to compare the various levels of information available—from the Initial Screener, through the Follow-up Screener, the Comprehensive Survey Instrument ,the DES medical examination and diagnostic procedures, including MER, and the functional analyses. Having all these data available on a large sample of individuals with moderate to severe disabilities, we will be able to analyze the adequacy of various levels of self-report information in approximating a valid and cost efficient disability determination.

Table 15. Cross-sectional Estimates of the Population with Disabilities Across Data Sources

Data	Year	Survey Questions	Population	Percent of Population with Disabilities
PSID[a]	1989	Do you have any nervous or physical condition that limits the type or the amount of work you can do? (Must have responded yes in both 1988 and 1989)	Aged 25 to 61	
			Men	9.2
			Women	10.6
CPS[b]	1990	Do you have a health problem or disability which prevents you from working or which limits the kind or the amount of work you can do? Or, Main reason did not work in 1989 was ill or disabled; or Current reason not looking for work is ill or disabled. (One period)	Aged 26 to 61	
			Men	8.1
			Women	7.8
SIPP[c]	1990	Do you have a physical, mental, or other health condition which limits the kind or amount of work you can do? (One period)	Aged 21 to 64	
			Men	11.7
			Women	11.6
SIPP[d]	1990	Do you have a physical, mental, or other health condition which limits the kind or amount of work you can do? (Must have responded yes in wave 3 and wave 6)	Aged 26 to 61	
			Men	9.8
			Women	9.8

[a] Panel Study of Income Dynamics (PSID) as reported in Burkhauser and Daly (1996b).
[b] Current Population Survey (CPS) as reported in Burkhauser and Daly (1996b).
[c] Survey of Income and Program Participation (SIPP) as reported in McNeil (1993).
[d] Survey of Income and Program Participation (SIPP) as reported in Burkhauser and Wittenberg (1996).

Source: Burkhauser and Daly (1996b), Burkhauser and Wittenburg (1996).

III. SSA Process Redesign Objectives

This section lays out the analysis and evaluation plans that pertain to the SSA process redesign objectives. As developed earlier, four such objectives have been specified by SSA for the DES. The four redesign objectives are:

1. Testing and validating the redesigned decisionmaking process as implemented by SSA;

2. Estimating the impact of the redesigned decisionmaking process on program size and costs;

3. Assessing changes in the number and kinds of persons who might be found disabled under the redesigned process, relative to the current process;

4. Identifying self-reported measures appropriate to estimating future program eligibility under the redesigned decisionmaking process.

Our plans are less definite for the redesign objectives and are, therefore, not developed in equivalent detail as for the four core objectives. This reflects the uncertainty over what the redesign will contain and when it will be implemented.

A. Redesign Objective 1 - Testing and Validating the Redesigned Decisionmaking Process

The purpose of the SSA process redesign project is to develop instrumentation and methodology for a new disability determination process (SSA, 1994a). A large part of the process of the redesign will be the development of alternative prototypes, each being evaluated to determine its feasibility and acceptability for administration on a broad scale throughout the United States. Ultimately, it is anticipated that one decision process will go forward for testing and validation. The first redesign research objective is to test and validate the new decision process.

For example, SSA is currently testing a single decisionmaker (SDM) model as a stand-alone process in several states. As we understand the model, a DDS examiner makes a decision as to eligibility for disability benefits, with a consulting physician or psychologist in an advisory role when called upon. In contrast, under the current system, a disability examiner and a physician or psychologist serve as co-adjudicators and jointly make disability determinations. Preliminary test results of the new system suggest that decisions under an SDM model may be more timely and more accurate. In addition, the new system saves administrative costs because medical consultants spend less time in the decisionmaking process.

We see little difficulty in adjusting our proposed activities to such changes in the decision methodology, assuming that the changes came before the commencement of disability decisionmaking for the main study. The major impact of the single decision maker model on the simulation process under the DES is that the "evidence" accumulated in the simulated case records would be given to the chosen DDS examiners for review, and not to a team of co-adjudicators. The DDS examiners would, at their discretion, ask for the advice of physicians or psychologists in deciding the case. In terms of testing the validity of the simulated process, the only potential problem that we foresee is a lack of data from the actual process for comparison. As above, the gold standard could provide the one means to evaluate the alternative procedures.

It is not known how much testing will have been done before the new instruments and decision processes are brought forward to the DES. Nevertheless, it is possible to speculate on the steps that will be necessary for further testing and validation. These steps include the following:

1. Testing instruments for use in the field;
2. Testing the content validity of instruments;
3. Developing methods for simulation;
4. Testing interrater reliability of the simulation process; and
5. Testing concurrent validity of the process.

Each of these steps is discussed below.

The redesign will likely include one or more instruments for screening individuals in and out of the disability determination process and for measuring the domains and dimensions most related to whether a person is able to perform SGA. We do not know how many instruments there will be; nor do we know what they will look like. Nevertheless, it will be necessary to ensure that they can be used as part of the DES process.

There are several types of validity. Face validity is based on a review of items by non-expert observers (Litwin, 1995). Content validity is an assessment of whether the measure being used to describe some real-world phenomenon seems to be describing that phenomenon. The assessment of content validity is a qualitative assessment and is usually done by obtaining feedback from "experts" on an instrument. "Experts" can mean medical, psychiatric, and physical therapy experts, but it can also mean persons with disabilities and their families. It is likely that face and content validity will have been assessed prior to the redesign being brought to the DES. If further assessment of content validity is necessary as part of the DES, we will obtain feedback from our expert consultants, as well as explore methods for testing instrumentation on persons with disabilities and their families.

Since the redesign process is still underway, there are as yet no policies, procedures, or SSA staff who have experience in the process. Therefore, an early step in testing and validating the redesign will be to develop policies and procedures and train appropriate individuals to become familiar with using the new process. As was proposed in the simulation of the current process, there will also have to be a stage for the development of the simulation of the new process. Many of the same issues are relevant (what information to use, whether the simulation process should be dynamic or static, who should make the determinations). Once these issues are resolved, staff will need to be trained both in the new decisionmaking process and the simulation protocol. The next step is to conduct the simulation and this can be done with "simulated" application folders given to trained decision makers, just as with the current process. (Core Objective 1)

Reliability is the extent to which a measuring device produces the same results upon multiple applications to the same phenomenon (Litwin, 1995). It refers to obtaining the same results each time something is measured (test-retest), obtaining the same results over time in the same individual (intrarater reliability), and obtaining the same results if different people use an instrument or process (interrater reliability). Assessment of test-retest and intrarater reliability require the passage of time to assess. Therefore, as in the simulation, we will only assess interrater reliability of the simulation of the redesign process. To do so we will use the same design and methodology that we propose for the current process simulation.

> One of the realities of the current method for determining disability is inconsistent decisionmaking. Allowance rates in different DDS offices range from a low of about 26 percent to a high of approximately 49 percent. Although it is possible that such variation is a partial result of geographic variation of impairments, it is also recognized that such differences may be the result of different decisionmaking procedures and policy interpretations. We consider good agreement to be Kappa values greater than 0.75 and poor agreement as less than 0.45 (Fleiss, 1981).

Since the goal of the redesign project is to develop a disability determination process that is "better, faster, and cheaper," a better process would be more reliable than the current process. A comparison of interrater reliability between simulations of disability determination in the current and redesigned systems can provide such information.

Since no gold standard to identify true disability currently exists, it has not been possible to determine whether the current disability determination process is valid (i.e., whether it is identifying all cases of true disability and the false positive rate is small). However, one of the goals of the DES is to develop a gold standard. Doing so will make it possible to compare the results of the current and redesigned process with this gold standard and determine the validity of the two different processes for determining disability. Such knowledge may also provide the American public with greater confidence in DDS and SSA decisions and may even assist in reducing the number of appeals from disability determinations. We will conduct a cost function assessment and quantify the dollar cost of false negatives and false positives in order to provide SSA with input for final policy decisions.

Concurrent validity refers to the fact that an instrument is measuring what it is supposed to be measuring (Litwin, 1995). To assess concurrent validity, it is usually necessary to compare the results of the instrument with the results of a gold standard. Since gold standards are rare, concurrent validity is difficult to test. However, one of the objectives of the DES is to develop a gold standard for disability decisionmaking. As a result, it will be possible to test the redesigned process against the gold standard to determine whether it has concurrent validity (i.e., is truly measuring disability).

To assess the concurrent validity of the redesigned method, we will compare the results obtained in the redesign simulation with the gold standard by using a standard 2 X 2 table (Table 16) and calculating sensitivity, specificity, false positive and false negative rates, predictive value, and accompanying confidence intervals (Fleiss, 1981; Gastwirth, 1987).

Table 16. Comparison of Results of Simulated Redesign Process with Gold Standard

Redesign Simulation	Gold Standard		
	Allowed/Disabled	Denied/Non-disabled	Total
Allowed/Disabled	a	B	a + b
Denied/Non-disabled	c	D	c + d
Total	a + c	b + d	a + c + b + d

Sensitivity in this case is the percentage of allowed claims which have been accurately classified as disabled [i.e., a/(a + c)] (MacMahon and Pugh, 1970). Specificity is defined as the percentage of denied claims which have been accurately classified as non-disabled [i.e., d/ (b + d)]. The predictive value of allowed claims is manifested by the probability that persons who are found disabled in the redesign simulation are truly disabled [a/(a + b)]. Conversely, the predictive value of denied claims is the probability that persons whose claim was found to be non-disabled in the simulation are truly non-disabled [d/(c + d)].

B. Redesign Objective 2 - Estimating the Impact of the Redesigned Decisionmaking Process on Program Size and Costs

While it is useful to have an estimate of the population of persons potentially eligible to receive disability benefits based on today's outlook, SSA is seeking an estimate of the impact on caseload and costs if various changes were made to the decisionmaking process. This objective calls for determining what the effect might be on future program size and cost. This section describes the modeling approach we will take to estimating the impact of the new decisionmaking process on program size and cost. It echoes the methodology presented earlier to accomplish core objective 2 (Section II-B). We begin with our general approach, summarizing the three steps to that approach; then we discuss the variables we will use.

Our general approach consists of three steps. In the first step, we will determine who applies for disability benefits. The second step will consist of determining who is eligible for benefits. Once we have determined who will apply and who will be granted benefits under the redesign process (and any other specified changes in policy or procedures), the stage is set for step three—a determination of the cost impacts of the redesign. This third step is relatively straightforward, but cannot be accomplished without conducting steps one and two.

We will use two basic approaches to evaluate the impact of potential changes in the SSA decisionmaking process. The difference in approaches is that one is driven by reduced form equations while the other is based on structural equations. The reduced form approach is relatively well established and will definitely yield usable projections. The structural equation approach is more sophisticated and offers the potential for greater understanding if it can be applied successfully. (See Section II-B)

In either event, the results of our first modeling step will be expressed as the number of individuals who are expected to apply for SSDI and SSI benefits under different sets of assumptions. These sets of assumptions (or alternative states) include individual, programmatic, and socioeconomic variables. It should be repeated that this approach assumes access to SSA administrative data for individuals in the DES sample who have given permission for these data to be released. Given structural and reduced form models that estimate the number of individuals who will apply for SSDI and SSI benefits as a function of various individual, programmatic, and socioeconomic variables, simulations of alternative states of the world can easily be constructed. Essentially, the estimated equations can be used as prediction equations by simply changing specific values or program parameters.

Our models will include policy variables and economic, health, and workplace factors (see table 9). The key policy variables with respect to the SSA disability programs relate to three dimensions of the system: (1) how benefits change over time; (2) how the disability determination process changes over time; and (3) how an individual's future wage earnings influence his or her future benefit level.

To assess how benefits change over time, we will develop an algorithm to estimate primary insurance amount for any feasible earnings history. Benefit levels influence individual decisions to apply for SSDI or SSI, other things being equal. The variable representing eligibility for disability benefits measures a person's perception of the likelihood of receiving benefits. As discussed earlier, the disability determination process takes place in several stages, including an appeal process. Burkhauser, Butler, Kim and Weathers (1997) used the state acceptance rates to identify their model. These rates vary across states and over time, and individuals are more likely to apply for benefits when the acceptance rate is higher.

Higher future wages increase the opportunity cost of applying for disability benefits, and thus individuals with higher expected wages are more likely to continue to work and postpone application to SSDI or SSI. Higher future wages also increase the future benefit amount and thus create a further incentive to delay application for disability benefits. Persons turned down for disability benefits can return to work, but Bound (1989) has shown that workers who have been denied benefits and return to work receive much lower earnings than before they applied for benefits. Therefore, an individual must consider not only the loss of earnings from not working during the application process, but also the likely lower earnings if that person's application is denied and he or she must return to work.

There are many other variables that may be influential in determining whether an individual with a given level of impairment applies for SSDI or SSI benefits. These variables were thoroughly presented in Section II-B above. Among the ones we will analyze here are:

1. State allowance rate;
2. State unemployment rate;
3. Employer accommodation possibilities;
4. Specific job characteristics;
5. Industry of last employment;
6. Union status at onset;
7. Employer size at onset;

8. Health condition at onset;

9. Education;

10. Tenure on the job;

11. Age at onset;

12. Marital status at onset; and

13. Income from other programs.

This last item is worthy of additional comment. Changes in state and federal income support programs may affect the number of people applying for SSDI and SSI benefits (Bound, Kossoudji, and Ricart-Moes, in preparation). In recent years, many states have eliminated or drastically reduced general assistance programs. As a result, individuals with disabilities who relied on these programs for income support may be forced to apply for SSDI or SSI benefits. Similarly, changes in the welfare system that limit the amount of time an individual can collect income support may prompt welfare recipients to apply for disability benefits. Thus, the linkage to other income maintenance programs is an important part of the analysis of applications to disability programs.

Summary of 3-Step Approach

As noted in Section II-B, the decision to apply for benefits is a complex one. We will use the same methodologies described in that section to estimate the time to disability application following the point at which a health condition first begins to limit one's ability to work, or in other words, whether an individual applies for disability benefits given a specific impairment (using a Hazard Model, Option Value Model, and Dynamic Programming Model).

The second step in our approach is to model the SSA disability determination process (i.e., who is awarded benefits, given that they applied). The SSA disability determination process will be modeled separately for SSDI and SSI, using data collected during the DES, as well as SSA administrative data. By relating the claimant success rate to various individual and program characteristics that will be collected during the DES, it also becomes possible to model the stages of disability determination with structural and reduced form models. A reduced form model derived from these results will enable simulation of alternative states of the world as shaped by administrative issues (e.g., the redesigned disability determination process). Thus, the output of this second step will be projections of the number of individuals granted SSDI and SSI benefits under alternative administrative procedures. We will also attempt to estimate structural models for the disability determination process.

The third step in this process will be to estimate the cost implications of the estimated SSDI and SSI populations. We will develop a costing model by generating annual and expected lifetime costs for individuals with particular characteristics. Based on the projected distribution of individual beneficiary characteristics (e.g., age, gender, earnings history), we will be able to obtain total system costs for SSDI and SSI through microsimulation techniques. The advantage of microsimulation is that it represents the full distribution of outcomes. Therefore, in a simulation of outcomes under a redesigned disability determination process, the number and characteristics of "winners" and "losers," as well as the distribution of amounts gained and lost, can be determined (Redesign Objective 3).

In addition to program benefit costs, we will work with SSA to develop an understanding of administrative costs. These costs can also be estimated as a function of the underlying applicant population and then can be simulated under alternative assumptions about that population. The process redesign presents additional analytic challenges. We will work with SSA to understand the ways in which process redesign will impact SSA administrative costs, for example, the single decision maker model. These costs can then be estimated under alternative states of the world, using the models developed earlier.

The process described above can be used to determine the impact on program size and cost of any changes in SSA legislation or policy. The proposed methodology can be used, in particular, to analyze the implications of the SSA disability process redesign. The SSA is expecting the redesign to have significant impacts on quality of service to claimants and savings in administrative costs (SSA, 1994a). There may be other unintended effects and these also need to be carefully reviewed. First, given the distribution of findings from the medical examination or functional analysis components of the DES, it would be possible to interpret assumptions about the ease or difficulty of entry into the SSDI or SSI systems as movements along the continuum of impairments. This in turn would produce alternative predictions of the number of persons qualifying for benefits. In addition, reducing the disability determination processing time may have an impact on the number of individuals applying for benefits. This is based on the behavioral assumption that future discounted benefits are one of the influences on application behavior. If the processing time is reduced, the future benefit stream increases slightly in value for that reason alone.

Based on the rich set of data from the DES medical exams, it would also be possible to compare different diagnostic procedures and/or other treatments utilized in different locations or with different subjects. Given informed assumptions about the magnitude of impact of such procedures (these could not be developed within our models, but could be generated by other means and then used as inputs to our models), we will also be able to model the implications for the number of beneficiaries and/or system costs of such changes.

We will work closely with SSA to ascertain the impacts of the redesign process as it affects both the likelihood of individuals to apply for benefits, and the probability that they will be granted benefits. Further, we will be able to do this for a large variety of alternative situations. Then, the models proposed in Section II-B can be effectively used to explore the cost implications of any changes that can be expressed in terms of their effect on applications or disability determinations.

C. Redesign Objective 3 - Assessing Changes in the Number and Kinds of Persons Who Would be Found Disabled Under the New Process, Relative to the Current Process

Using a simulation approach to estimate program size, instead of a modeling approach, it will be relatively straightforward to examine the changes in the number and kinds of persons who would be found disabled under the new process, relative to the current process. This approach consists of comparing the results of the redesign of disability determination simulation with the results of the simulation under the current process. Such a comparison will allow us to determine the prevalence of disability as judged by the current decisionmaking process and the prevalence of disability according to the redesigned process. In addition, we will be able to identify the characteristics of those individuals who were found disabled in one decisionmaking process but not the other.

We therefore propose a straightforward analysis which consists of developing a table similar to Table 17 and calculating estimates of the number of individuals in the U.S. population, and confidence intervals surrounding those estimates.

Table 17. Comparison of Results of Simulated Redesign Process with Current Process

Redesign Simulation	Current Process Simulation		
	Disabled	Non-disabled	Total
Disabled	a	b	a + b
Non-disabled	c	d	c + d
Total	a + c	b + d	a + c + b + d

There are two important prevalence statistics—the prevalence of disability as judged by each disability determination process. The prevalence of disability using the current system simulation can be calculated as (a + c)/(a + c + b + d). The prevalence of disability using the redesign is (a + b)/(a + c + b + d). It will be important to examine the difference in prevalence from the two decisionmaking processes as well as provide an estimate of the numbers in the U.S. population and accompanying confidence intervals. We will use WesVarPC to calculate weighted estimates and confidence intervals.

SSA also wants to know about the similarities and differences between the current and redesign systems. Four comparison groups will provide such information:

1. Those found disabled in both the current and redesign process
[(a / (a + c + b + d)];

2. Those found non-disabled in both the current and redesign process
[d/ (a + c + b + d)];

3. Those found disabled in the current process, but not in the redesign process
[c/(a + c + b + d)]; and

4. Those found disabled in the redesign process, but not in the current process
[b/(a + c + b + d)].

All the data of the DES (see table 9) will be available for these individuals, plus the matched SSA administrative data for most. This will enable a careful analysis of “winners” and “losers” under the new disability determination system.

Thus, we will examine the discordant pairs (b and c—those who were found disabled in the redesign simulation but not in the current system simulation and those not found disabled in the redesign simulation but found disabled in the current system simulation). These groups will be analyzed by age, gender, type of impairment (e.g., mental vs. physical, musculoskeletal, respiratory), and other relevant variables to determine the number and kinds of persons who would be found disabled under the new process relative to the current process. Analysis will consist of simple cross-tabulations. Although we may wish to provide population estimates using WesVarPC, cell size may be small for some breakdowns. Therefore, precision of estimates may be low.

D. Redesign Objective 4 - Identifying Self-Reported Measures Appropriate to Estimating Program Eligibility Under the New Process

In Section II-D above, we described our methodology for identifying a subset of self-reported measures appropriate to estimating program eligibility for the current process. The process included three steps: (1) developing estimation models for predicting eligibility from self-reported data; (2) determining the sensitivity and specificity of predictive models; and (3) identifying ongoing surveys which provide similar self-reported data to enable tracking of future trends in the eligible population. The three steps are the same in meeting this objective for the redesign.

However, because the redesign would use different kinds of information in a different decisionmaking process, there will be differences in the variables chosen to include in the models and, of course, in the influence of those variables. For example, the redesign may substitute an Index of Disabling Impairments (Index) for the medical Listings of Impairments (Listings). Although the Index is yet to be developed, it is anticipated that it will contain a shorter list of impairments. Whatever the contents of the Index, the self-reported measures chosen for inclusion in the predictive equations will need to reflect the impairments and functional abilities contained in the Index. Moreover, the steps that will be used in the redesign need to be reflected in the statistical modeling process. The process of estimation will be the same, however, as would the process for determining sensitivity and specificity and developing recommendations for augmenting future surveys.

Using these analytical techniques and the database developed in the DES, as augmented by SSA administrative records, the project will assist SSA in redesigning the disability determination process to be fairer, faster, and cheaper to administer.

References

Aarts, Leo J. M., Richard V. Burkhauser, and Phillip deJong, (eds.). 1996. *Curing the Dutch Disease: An International Perspective on Disability Policy Reform.* Aldershot, Great Britain: Avebury.

Aarts, Leo J., and Philip R. deJong. 1990. *Economic Aspects of Disability Behavior*. Amsterdam: North-Holland.

Acs, Gregory, and Pamela Loprest. 1999. "The Effect of Disabilities on Exits from AFDC." J*ournal of Policy Analysis and Management.* 18:1, 28-49.

Ausink, John A., and David A. Wise. 1993. "The Military Pension, Compensation, and Retirement of U.S. Air Force Pilots." NBER Working Paper No. 4593, December.

Bane, Mary Jo, and David T. Ellwood. 1994. *Welfare Realities: From Rhetoric to Reform.* Cambridge and London: Harvard University Press.

Barlow, W., M. Lai, and S. P. Azen. 1991. "A Comparison of Methods for Calculating a Stratified Kappa." *Statistics in Medicine.* 10:1465-72.

Benítez-Silva, Hugo, Moshe Buchinsky, Hiu Man Chan, John Rust, and Sofia Sheidvasser. 1999. "An Empirical Analysis of the Social Security Disability Application, Appeal, and Award Process." *Labour Economics*. 6:2, 147-78.

Berkovec, James, and Steven Stern. 1991. "Job Exit Behavior of Older Men." *Econometrica.* 59:1: 189-210.

Bird, Chloe, and Allen Freemont. 1991. "Gender, Time Use and Health." Journal of Health and Social Behavior. 32:114-129.

Blank, Rebecca. 1997. *It Takes a Nation: A New Agenda for Fighting Poverty.* New York: Russell Sage Foundation.

Blank, Rebecca. 1997. "What Causes Public Assistance Caseloads to Grow?" National Bureau of Economic Research Working Paper No. 5149.

Blau, David M. 1994. "Labor Force Dynamics of Older Men." *Econometrica.* 62:117-56.

Bound, J. (1989). "The Health and Earnings of Rejected Disability Insurance Applicants." *American Economic Review.* 79:482-503.

Bound, John. 1991. "Self-Reported Versus Objective Measures of Health in Retirement Models." *Journal of Human Resources.* 26:l, 106-38.

Bound, John, S. Kossoudji, and G. Ricart-Moes. 1998. "The Ending of General Assistance and SSI Disability Growth in Michigan: A Case Study." In Kalman Rupp and David Stapleton, (eds.), *Growth in Cash Benefits for Disability: Explanations and Policy Implications.* Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone. 1984. *Classification and Regression Trees*. Wadsworth: Belmont, CA.

Brennan, R. L., and D. J. Prediger. (1981). Coefficient Kappa: Some Uses, Misuses, and Alternatives." *Educational and Psychological Measurement.* 41:687-99.

Brown, C. C., and L. R. Muenz. 1976. "Reduced Mean Square Error Estimation in R x C Contingency Tables." *Journal of the American Statistical Association.* 71:176-82.

Bull, B., and A. Donner. 1987. "The Efficiency of Multinomial Logistic Regression Compared with Multiple Group Discriminant Analysis." *Journal of the American Statistical Association.* 82:1118-22.

Burbridge, John B. and Alan L. Robb. 1980. "Pensions and Retirement Behavior." *Canadian Journal of Economics.* 13:3, 421-37.

Burkhauser, Richard V., J. S. Butler, and Yang-Woo Kim. 1995. "The Importance of Employer Accommodation on the Job Duration of Workers with Disabilities: A Hazard Model Approach." *Labour Economics.* 2:109-30.

Burkhauser, Richard V., J. S. Butler, Yang-Woo Kim, and Robert Weathers. 1997. "The Importance of Accommodation on the Timing of Disability Insurance Application: A Choice-Based Semiparametric Hazard Model." Aging Studies Program Paper No. 12, Center for Policy Research, The Maxwell School. Syracuse, NY: Syracuse University, October.

Burkhauser, Richard V., and Mary C. Daly. 1997. "Employment and Economic Well-Being Following the Onset of a Disability: The Role for Public Policy." In Jerry L. Mashaw, Virginia Reno, Richard V. Burkhauser, and Monroe Berkowitz, (eds.), *Disability, Work and Cash Benefits.* Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Burkhauser Richard V., and David C. Wittenburg. 1996. "The Changing Face of the Population with Disabilities in the 1990s: An Analysis Using SIPP Data." Paper presented at the NIDRR/SSA Research and Policy Education Workgroup, Detroit, May 21, 1996.

Butler, J. S., and Robert Moffitt. 1982. "A Computationally Efficient Quadrature Procedure for the One-Factor Multinomial Probit Model." *Econometrica.* 50:761-64.

Bye, Barry V., and Evan S. Scheckter. 1982. "A Latent Markov Model Approach to the Estimation of Response Errors in Multiwave Panel Data." *Journal of the American Statistical Association.* 81:394, 375-80 .

Campbell, G. 1994. "Advances in Statistical Methodology for the Evaluation of Diagnostic and Laboratory Tests." *Statistics in Medicine.* 13:499-508.

Choi, Namkee. 1998. "A Comparative Study of Elderly SSI Recipients, Denied Applicants and Eligible Nonapplicants." *Journal of Aging and Social Policy.* 10:2, 7-28.

Coe, Richard D. 1985. Nonparticipation in the SSI Program by the Eligible Elderly. *Southern Economics Journal.* 51:3, 891-97.

Danziger, Sandra, Mary Corcoran, Sheldon Danziger, Colleen Heflin, Ariel Kalil, Judith Levine, Daniel Rosen, Kristin Seefeldt, Kristine Siefert and Richard Tolman. 1999. "Barriers to the Employment of Welfare Recipients", Working Paper, Poverty Research and Training Center, University of Michigan.

Daula, Thomas, and Robert Moffitt. 1991. "Estimating a Dynamic Programming Model of Army Reenlistment Behavior." In C. Gilroy, D. Home, and D. Smith (eds.), M*ilitary Compensation and Personnel Retention: Models and Evidence.* Alexandria, VA: Army Research Institute.

Daula, Thomas, and Robert Moffitt. 1995. "Estimating Dynamic Models of Quit Behavior: The Case of Military Reenlistment." *Journal of Labor Economics.* 13:499-523.

Donner, Allan, and Michael Eliasziw. 1987. "Sample Size Requirements for Reliability Studies." *Statistics in Medicine*. 6: 441-48.

Dunn, G. (1992). "Design and Analysis of Reliability Studies" *Statistical Methods in Medical Research.* 1:123-57.

Efron, B., and R. J. Tibshirani. 1993. *An Introduction to the Bootstrap*. Chapman and Hall: London.

Executive Office of the President of the Unites States, Council of Economic Advisers. 1999. "The Effects of Welfare Policy and the Economic Expansion on Welfare Caseloads: An Update."

Fields, Gary, and Olivia Mitchell. 1984. "Economic Determinants of the Optimal Retirement Age--An Empirical Investigation." *Journal of Human Resources.* 19:245-62.

Fishburn, Peter C. "Foundations of Risk Management." 1988. *Encyclopedia of Statistical Sciences,* Vol. 8. New York: John Wiley and Sons.

Fleiss, J. L. 1981. *Statistical Methods for Rates and Proportions, 2nd Edition.* New York: Wiley.

Freedman, Vicki, and Douglas Wolf. 1995. "A Case Study in Multiple Imputation." *Demography.* 32:459-70.

Gastwirth, J. L. 1987. "The Statistical Precision of Medical Screening Procedures: Application to Polygraph and AIDS Antibodies Test Data" (discussion: pp.223-228). *Statistical Science.* 2:213-22.

Gong, G. (1986). "Cross-validation, the Jackknife, and the Bootstrap: Excess Error Estimation in Forward Logistic Regression." *Journal of the American Statistical Association.* 81:108-13.

Graham, P., and R. Jackson. 1993. "The Analysis of Ordinal Agreement Data: Beyond Weighted Kappa." *Journal of Clinical Epidemiology.* 46:1055-62.

Greenblum, Joseph, and Barry Bye. 1987. "Work Values of Disabled Beneficiaries." *Social Security Bulletin.* 50:4, 67-74.

Halpern, Janice, and Jerry A. Hausman. 1986. "Choice under Uncertainty: A Model of Applications for the Social Security Disability Insurance Program." *Journal of Public Economics.* 31:131-61.

Härdle, Wolfgang, and Thomas M. Stoker. 1989. "Investigating Smooth Multiple Regression by the Method of Average Derivatives." *Journal of the American Statistical Association.* 84:986-95.

Harris, Kathleen. 1996. "Life After Welfare: Women, Work and Repeat Dependency." *American Sociological Review*. 61:407-26.

Haveman, Robert and Barbara Wolfe. "The Economics of Disability and Disability Policy", *Handbook of Health Economics.* Elsevier, Amsterdam, forthcoming.

Hennessey, John C. 1996. "Job Patterns of Disabled Beneficiaries." *Social Security Bulletin.* 59:4, 4.

Hennessey, John C. 1997. "Factors Affecting the Work Efforts of Disabled-Worker Beneficiaries." *Social Security Bulletin.* 60:3, 3-20.

Hennessey, John C., and J. Dykacz. 1989. "Projected Outcomes and Length of Time in the Disability Insurance Program." *Social Security Bulletin.* 52:9, 2-41.

Hill, Daniel H. 1990. "An Endogenously-Switching Ordered-Response Model of Information, Eligibility and Participation in SSI." *The Review of Economics and Statistics.* 72:2, 368-71.

Holzer, Harold. 1996. What Employers Want: Job Prospects for Less-Educated Workers. New York: Russell Sage Foundation.

Holzer, Harry J. 1999. "Will Employers Hire Welfare Recipients? Recent Survey Evidence from Michigan." *Journal of Policy Analysis and Management*. 18:3.

Hoynes, Hillary, and Robert Moffitt. 1996. "The Effectiveness of Financial Work Incentives in Social Security Disability Insurance and Supplemental Security Income: Lessons from Other Transfer Programs." In Jerry L. Mashaw, Virginia Reno, Richard Burkhauser, and Monroe Berkowitz, (eds.), *Disability, Work, and Cash Benefits.* Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Hunt, H. Allan, R. V. Habeck, P. Owens, and D. Vandergoot. 1996. "Disability and Work: Lessons from the Private Sector." In Jerry L. Mashaw, Virginia Reno, Richard Burkhauser, and Monroe Berkowitz, (eds.), *Disability, Work and Cash Benefits,* Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Ihlanfeldt, Keith and Lars Sjoquist. 1998. "The Spatial Mismatch Hypothesis: A Review of Recent Studies and Their Implications for Welfare Reform." Housing Policy Debate. 9:4, 849-92.

Juster, F. Thomas, and Richard Suzman. 1995. "An Overview of the Health and Retirement Study." In Richard V. Burkhauser and Paul J. Gertler, (eds.), *Special Issue, Journal of Human Resources,* Vol. 30 (December): S7-S56.

Kain, John. 1992. "The Spatial Mismatch Hypothesis: Three Decades Later." *Housing Policy Debate*. 3:2, 371-460.

Kallenberg, W. C. 1989. "Estimation and Testing in Large Binary Contingency Tables." *Journal of Multivariate Analysis.* 30:205-26.

Keane, Michael P., and Kenneth I. Wolpin. 1994. "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence." *Review of Economics and Statistics*. 76:4, 648-72.

Kramer, Michael S. (1988). *Clinical Epidemiology and Biostatistics.* Springer-Verlag Berlin Heidelberg.

Kreider, Brent. 1998. "Workers' Applications to Social Insurance Programs When Earnings and Eligibility Are Uncertain." *Journal of Labor Economics.* 16:4, 848-77.

Lahiri, Kajal, Denton R. Vaughan, and Bernard Wixon. 1995. "Modeling SSA's Sequential Disability Determination Process Using Matched SIPP Data." *Social Security Bulletin.* 58:4, 3-42.

Landis, J. R., and G. G. Koch. 1977. "The Measurement of Interrater Agreement for Categorical Data." *Biometrics.* 33:159-74.

Linton, S. 1998. *Claiming Disability*. New York: New York University Press.

Lipsitz, S. R., N. M. Laird, and D. P. Harrington. 1990. Finding the Design Matrix for the Marginal Homogeneity Model." *Biometrika.* 77:353-58.

Litwin, Mark S. 1995. *How to Measure Survey Reliability and Validity*. London: SAGE Publications.

Lucas, Robert E., Jr. 1976. "Econometric Policy Evaluation: A Critique." In Karl Brunner and Allan H. Meltzer (eds.), *The Phillips Curve and Labor Markets,* Vol. 1, Camegie-Rochester Conference Series on Public Policy. Amsterdam: North Holland Publishing Company, p. 1946.

Lumsdaine, Robin L., James H. Stock, and David A. Wise. 1990. "Three Models of Retirement: Computational Complexity versus Predictive Validity." NBER Working Paper No. 3558, December.

Lumsdaine, Robin L., James H. Stock, and David A. Wise. 1992. "Pension Plan Provisions and Retirement: Men & Women, Medicare, and Models." NBER Working Paper No. 4201, December.

MacMahon, Brian, and Thomas F. Pugh. 1970. *Epidemiology*. *Principles and Methods.* Boston: Little, Brown and Company.

Manski, Charles, and S. Lerman. 1977. "The Estimation of Choice Probabilities from Choice Based Samples." *Econometrica*. 45:8, 78-88.

Mashaw, Jerry L., and Virginia P. Reno, (eds). 1996. "Balancing Security and Opportunity: The Challenge of Disability Income Policy." Disability Policy Panel Interim Report. National Academy of Social Insurance, Washington, D.C. 102-106.

Mashaw, Jerry L., Virginia P. Reno, Richard V. Burkhauser, and Monroe Berkowitz, (eds). 1996. *Disability, Work, and Cash Benefits.* Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

MathSoft Software. 1997. S-Plus 4 Guide of Statistics. Data Analysis Products Division, MathSoft. Seattle.

McNally, James, and Douglas A. Wolf. 1996. "Family Structure and Institutionalization: Results from Merged Microdata." Papers in Microsimulation Series Paper No. 2, Center for Policy Research, The Maxwell School. Syracuse, NY: Syracuse University, January.

Messenger, R. C., and M. L. Mandell. 1972. "A Model Search Technique for Predictive Nominal Scale Multivariate Analysis." *Journal of the American Statistical Association.* 67:768-72.

Muller, L. Scott. 1992. "Disability Beneficiaries Who Work and Their Experience Under Program Work Incentives." *Social Security Bulletin.* 55:2, 2-19.

Muller, L. Scott, Charles G. Scott, and Barry V. Bye. 1996. "Labor-Force Participation and Earnings of SSI Disability Recipients: A Pooled Cross-Sectional Times Series Approach to the Behavior of Individuals." *Social Security Bulletin.* 59:1.

National Emergency TANF Datafile as of 12/9/98, Table 30.

Oi, Walter Y. 1996. "Employment and Benefits for People With Diverse Disabilities." In Jerry Mashaw, et. al., (eds.), *Disability, Work, and Cash Benefits*. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, 103-28.

Olson, Krista, and LaDonna Pavetti. 1996. *Personal and Family Challenges to the Successful Transition from Welfare to Work.* Washington, D.C.: The Urban Institute.

Parsons, Donald O. 1980. "The Decline in Male Labor Force Participation." *Journal of Political Economy.* 88:l, 117-34.

Parsons, Donald O. 1991a. "The Health and Earnings of Rejected Disability Insurance Applicants: Comment." *American Economic Review.* 81:5, 1419-26.

Parsons, Donald O. 199lb. "Self-Screening in Targeted Public Transfer Programs." *Journal of Political Economy.* 99:859-76.

Parsons, Donald O. 1996. Imperfect 'Tagging' in Social Insurance Programs. *Journal of Public Economics.* 62:183-207.

Rao, C. R. 1973. *Linear Statistical Inference.* New York: Wiley.

Robins, J., N. Breslow, and S. Greenland. 1986. "Estimators of the Mantel-Haenszel Variance Consistent in Both Sparse Data and Large-Strata Limiting Models." *Biometrics.* 42:311-23.

Rupp, Kalman, and Charles Scott. 1995. "Length of Stay in the Supplemental Security Income Disability Program." *Social Security Bulletin.* 58:1, 29-47.

Rupp, Kalman, and Charles Scott. 1996. "Trends in the Characteristics of DI and SSI Disability Awardees and Duration of Program Participation." *Social Security Bulletin.* 59:1, 3-21.

Rupp, Kalman, and David Stapleton. 1995. "Determinants of the Growth in the Social Security Administration's Disability Programs—An Overview." *Social Security Bulletin.* 58:4, 43-70.

Rupp, Kalman, and David Stapleton, (eds.). 1998. *Growth in Disability Benefits.* Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Rust, John. 1989. "A Dynamic Programming Model of Retirement Behavior." In D. Wise (ed.), *The Economics of Aging*. Chicago: University of Chicago Press, 359-398.

SAS Institute Inc. 1997. SAS/STAT© Software: Changes and Enhancements through Release 6.12, Cary, NC: SAS Institute.

Sauermann, J. 1989. "Bootstrapping the Maximum Likelihood Estimator in High-Dimensional Log-linear Models." *Annals of Statistics.* 17:1198-1216.

Scotch, R. K., and K. Schriner. 1997. "Disability and Human Variation: Implications for Policy." *Annals, AAPSS.* 549:148-60.

Scott, Charles G. 1992. "Disabled SSI Recipients Who Work." *Social Security Bulletin.* 55:1, 26-36.

Social Security Disability Insurance: Multiple Factors Affect Beneficiaries' Ability to Return to Work (Letter Report, 01/12/98, GAO/HEHS-98-39).

Simonoff, J. S. 1983. "A Penalty Function Approach to Smoothing Large Sparse Contingency Tables." *Annals of Statistics.* 11:208-18.

Stem, Steven. 1989. "Measuring the Effect of Disability and Labor Force Participation." *Journal of Human Resources.* 24:3, 361-95.

Stock, James H., and David A. Wise. 1990. "Pensions, the Option Value of Work, and Retirement." *Econometrica.* 58:1151-80.

Stoker, Thomas M. 1992. Lectures on Semiparametric Economics, Core Lecture Series, Core Foundation, Louvain-la-Neuve, Universityé Catholique de Louvain.

Turner, Margery, Michael Fix, and Raymind Struyk. 1991. "Opportunities Denied, Opportunities Diminished: Racial Discrimination in Hiring." Urban Institute Report 91-9.

U.S. Department of Health and Human Services, Social Security Administration. 1994a. *Plan for a New Disability Claim Process.* SSA Pub. No. 01-005.

U.S. Department of Health and Human Services, Social Security Administration. 1994b. *Disability Process Redesign: Next Steps in Implementation* SSA Pub. No. 01-006.

U.S. Department of Health and Human Services, Social Security Administration. 1995. *Annual Statistical Supplement.* Washington, DC: U.S. Government Printing Office.

U.S. Department of Health and Human Services. 1999. Second Annual Report to Congress. Appendix 14:1 (August).

U.S. Department of Health and Human Services. 1999. Temporary Assistance for Needy Families Program, Second Annual Report to Congress, p. 7.

U.S. General Accounting Office. 1996. Testimony before the Special Committee on Aging, U.S. Senate. "Social Security: Disability Programs Lag in Promoting Return to Work." T-HEHS-96-147.

U.S. General Accounting Office. 1997. Testimony before the Subcommittee on Social Security, Committee on Ways and Means, House of Representatives. "Social Security Disability: Improving Return-to-Work Outcomes Important, but Trade-offs and Challenges Exist." T-HEHS-97-186.

U.S. House of Representatives, Committee on Ways and Means. 1996. *1996 Green Book*. Washington, D.C.: Government Printing Office.

Velema, J. P., M. Blettner, M. Restrepo, and N. Munoz. 1991. "The Evaluation of Agreement by Means of Log-linear Models: Proxy Interviews on Reproductive History Among Floriculture Workers in Colombia." *Epidemiology.* 2:107-15.

Verbrugge, L. M., and A. M. Jette. 1994. "The Disablement Process." *Social Science Medicine*." 38:1, 1-14.

Wolf, Douglas. 1997a. "Microsimulation as a Tool for Population Projection." Presented at the annual meetings of the Population Association of America, Washington, DC, March 27-29, 1997.

Wolf, Douglas. 1997b. "Technical Issues in the Design and Implementation of Dynamic Microsimulation Models." In W. Lutz (ed.), *FAMSIM-Austria: A Prototype Dynamic Microsimulation Model for the Evaluation of Family Policies and Projections*, Chapter 4. To be published by the Austrian Institute for Family Studies.

Wolf, Douglas, and Sarah Laditka. 1996. "Stochastic Modeling of Active Life and its Expectancy." Presented at the REVES9 conference, Rome, December 11-13, 1996.

Wolf, Douglas A., Jan Ondrich, Kenneth G. Manton, Eric Stallard, Max A. Woodbury, and Larry Corder. 1995. "A Model for Simulating Life Histories of the Elderly: Model Design and Implementation Plans." Papers in Microsimulation Series Paper No. 1, Center for Policy Research, The Maxwell School. Syracuse, NY: Syracuse University, September.

Yelin, E. H., and P. P. Katz. 1994. "Making Work More Central to Work Disability Policy. *Milbank Quarterly.* 72:593-619.

Yelin, E. H. and P. P. Katz. 1994. "Labor Force Trends of Persons with and without Disabilities. *Monthly Labor Review.* 117:36-42.

Yelowitz, Aaron. 1997. "Why Did the SSI-Disabled Program Grow So Much? Disentangling the Effect of Medicaid." National Bureau of Economic Research, Working Paper 6139.

Ziliak, James P., David N. Figlio, Elizabeth E. Davis, and Laura S. Connelly. 1997. "Accounting for the Decline in AFDC Caseloads: Welfare Reform or Economic Growth?" Institute for Research on Poverty, University of Wisconsin-Madison, Discussion Paper 1151-97.

Appendix A

Structural Models

Option Value Model

We follow Lumsdaine, Stock, and Wise (1990) in setting up our model, which includes utility functions for labor earnings and SSDI benefits. We begin with the option value model. Note that many letters refer to time in various aspects. Let the current year be t. One can apply for disability benefits in year r; $r \geq t$. The subscript s refers to any time period. One will retire in the year in which one reaches the age of 65, regardless of other factors. One dies in year d. If one applies for disability benefits, the probability of being approved is a(t) and of going back to work if not approved is $\delta(t)$. Income while still working is Y_s, and income with SSDI benefits is D_s. Note that SSDI becomes retirement income at age 65, but we will continue to call it D. If one is turned down for SSDI, income is B_s if one does not return to work.

Utility associated with labor earnings is

$$U_w(Y_s) = (Y_s)^{\gamma} + \omega_s , \tag{1}$$

where γ is a parameter showing the decreasing ($\gamma < 1$) or increasing ($\gamma > 1$) marginal utility of income. Utility associated with nonlabor income is

$$V_o(Z_s) = k^{\gamma} Z_s^{\gamma} + \xi_s , \tag{2}$$

where Z is the income that results from disability application, and k is a parameter relating the utility value of wage income to the utility value of nonwage income.

The nonstochastic basis of disability application in the absence of random utility shocks ω_s or ξ_s is

$$V_t(r) = \sum_{s=t}^{r-1} \beta^{s-t} U_w(Y_s)$$

$$+ \sum_{s=r}^{d} \beta^{s-t} \left[\alpha(t) V_o(D_s(r)) + (1-\alpha(t))\delta(t) U_w(Y_s(r)) + (1-\alpha(t))(1-\delta(t)) V_o(B_s(r)) \right].$$

For ease of discussion of formulae from this point on, we will let $U_D(D_s)$ denote the weighted average in brackets, which is the expected value of applying for disability benefits. As in the Lumsdaine, Stock, and Wise (LSW) model, we assume that nonlabor income may produce more or less utility than labor income.

We assume that the change in utility is the same, regardless of which outcome (approval or rejection without a return to work) occurs.

In the specification of the disturbances, we will try a variation on the LSW model. They assume that $\omega_s = \rho\,\omega_{s-1} + \varepsilon_{\omega s}$ and $\xi_s = \rho\,\xi_{s-1} + \varepsilon_{\xi s}$. Note that the same *p* applies to both. Then they define $\nu_s = \omega_s - \xi_s$ and it follows that $\nu_s = \rho\,\nu_{s-1} + \varepsilon_s$. This specification has worked well in several papers on retirement, but we also propose to try an alternative, in which there is a random effect for each person, as in Berkovec and Stern (1991) and Daula and Moffitt (1995). Defining $\omega_s = \mu_\omega + \varepsilon_{\omega s}$ and $\xi_s = \mu_\xi + \varepsilon_{\xi s}$ to derive our error $\theta_s = \omega_s - \xi_s$ yields $\theta_s = \mu + \varepsilon_s$. Here μ is a random effect. In both the LSW model and the random effects model, the disturbances are assumed to be normally distributed.

The value of applying for disability benefits now (period *t*) is

$$V_t(t) = \sum_{s=t}^{d} \beta^{s-t}\, U_D(D_s(t)). \tag{4}$$

Define the expected value of applying for disability at some date in the future, *r*, minus the expected value of doing so now to be $G_t(r)$: $G_t(r) = E_t(V_t(r)) - E_t(V_t(t))$. Substituting in the above formulae leads to

$$G_t(r) = \sum_{s=t}^{r-1} \beta^{s-t}\, E_t\left(Y_s^{\gamma}\right) + \sum_{s=r}^{d} \text{betasup}\, s\text{-}t\; k^{\gamma} E_t(D_s(r)^{\gamma}) - \sum_{s=t}^{d} \beta^{s-t}\, k^{\gamma}\, E_t(D_s(t)^{\gamma}) + \sum_{s=t}^{r-1} \beta^{s-t}\, E_t(\nu_s).$$

Up to this point, the date of death is assumed fixed, but now we define the probability of living to period *s*, given that one has already lived to period *t*, *Pr*(live to s | already lived to t), as $\pi(s|t)$. Now,

$$G_t(r) = \left[\sum_{s=t}^{r-1} \beta^{s-t}\, \pi(s|t)\, E_t\left(Y_s^{\gamma}\right) + \sum_{s=r}^{d} \beta^{s-t}\, k^{\gamma}\, \pi(s|t)\, E_t(D_s(r)^{\gamma}) - \sum_{s=t}^{d} \beta^{s-t}\, k^{\gamma}\, \pi(s|t)\, E_t(D_s(t)^{\gamma}) \right] + \sum_{s=t}^{r-1} \beta^{s-t}\, \pi(s|t)\, E_t(\nu_s). \tag{6}$$

The term in brackets is $g_t(r)$, the systematic, exogenous portion of utility associated with applying for disability in year r. The last term is the stochastic portion of utility, $\varphi_t(r)$. In the LSW model, $E_t(\nu_s) = \rho^{s-t}\, \nu_t$, and

$$\phi_t(r) = \sum_{s=t}^{r-1} \beta^{s-t} \ \pi(s|t) \ \rho^{s-t} \ \nu_t = \nu_t \left[\ \sum_{s=t}^{r-1} \beta^{s-t} \ \pi(s|t) \, \rho^{s-t} \ \right] = \nu_t \ K_t(r). \tag{7}$$

In the random effects model, $E_t(\theta_s) = \mu + \varepsilon_t(\text{when s = t}), \text{or } \mu \text{ (otherwise)}$, and

$$\phi_t(r) = \varepsilon_t + \sum_{s=t}^{r-1} \beta^{s-t} \ \pi(s|t) \ \mu = \varepsilon subt + \mu \left[\sum_{s=t}^{r-1} \beta^{s-t} \pi(s|t) \right] = \varepsilon_t + \mu \ K_t(r), \text{and} \tag{8}$$

$$G_t(r) = g_t(r) + K_t(r) \ \mu + \varepsilon_t. \tag{9}$$

Note that $K_t(r)$ in (8) and (9) differs from $K_t(r)$ in (7) because this random effects specification has no auto-regressive structure and thus no r term. We allow for a period-specific disturbance in the random effects model that is absent in the LSW model. If the disturbances are all independent, then the period-specific disturbance is all there is. Thus, we believe our assumption is less restrictive since our random effects specification includes both the independent and LSW cases. The LSW model can produce independent disturbances when r is zero.

The LSW model leads to the following equations for the year in which one applies for disability benefits. One applies for disability benefits in year t if $g_t(s) + K_t(s)\,\mathrm{v}_t < 0$ for $t + 1 \le s \le d$. One applies for disability benefits in year $r > t$ if $g_r(s) + K_r(s)\,\mathrm{v}_r < 0$ for $r + 1 \le s \le d$ and for $t' < r$, there is an s such that $g_{t'}(s) + K_{t'}(s) \ \nu_{t'} > 0$. Thus, one must compute $g_i(j)$ and $K_i(j)\mathrm{v}_i$ for $t \le i < j \le d$. This step is recursive and entails Taylor expansion of Y^{γ} and $k^{\gamma} D_s^{\gamma}$ if a linear equation is desired.

If the random utility component in any future period is less than the negative of the maximum of $g_i(j)/K_i(j)$ over j, given i—defined as $g_i^* \ / \ K_i^*$, it is optimal to postpone the application for disability benefits to the future. Now define r as the year in which application for disability benefits occurs. We assume one must retire in year T, which we interpret to be the year one reaches age 65. To calculate the probabilities for a likelihood function or other estimation method, define the following:

$$Pr(r = t) = Pr \ (g_t^* \ / \ K_t^* < - \nu_t) \ ;$$

$$Pr(r = s > t) = Pr(g_t^* \ / \ K_t^* > - \nu_t, \ g_{t+1}^* \ / \ K_{t+1}^* > - \nu_{t+1}, \ldots, \ g_s^* \ / \ K_s^* < - \nu_s) \ ;$$

$$Pr(r > T) = Pr(g_t^* \ / \ K_t^* > - \nu_t, \ldots, \ g_T^* \ / \ K_T^* > - \nu_T^* \ ) \ ).$$

Note that all of these probabilities add to unity. They entail multivariate normal cumulative density functions. Stock and Wise (1990) assume AR(1) errors, so they have multivariate integrals to solve. If the alternative, random effects assumption is made, then only univariate integrals result, using the Butler and Moffitt (1982) simplification, which is also employed in the dynamic programming model of Daula and Moffitt (1995), below. This simplification loses one part of an argument of Stock and Wise (1990), where they state that "there are two equivalent ways to see that uncertainty about the future is reduced as the planning horizon is shortened, ... there are fewer future random components of utility to cumulate [this is still true] ... [and] the uncertainty about the value of future random effects is reduced [this is lost] ..." However, they require that the AR parameter be the same in both utility functions, while the alternative assumption imposes no requirements on the random effects in the two utility functions.

To summarize, the steps in the option value estimation are first to estimate g and K using Taylor expansions of the nonlinear functions and autoregressions to predict several different types of income, then to estimate the univariate or multivariate normal likelihood functions for retirement age. The latter step produces parameters of the utility function and the variances. These steps in principle could be combined to produce correct standard errors not assuming independence of the disturbances at the two levels of estimation (*g*/*K* and application for SSDI) using the alternative (random effects) assumption and Generalized Method of Moments (GMM).

Dynamic Programming Model

The steps in the estimation of the dynamic programming model are as follows. (1) Calculate the expected utility of retiring at age 65. (2) Calculate the expected maximum utility of applying for disability benefits at age 64, which is a maximum over two independent random variables or over two sums of a random effect and an independent disturbance. The probability of applying at age 64 is the probability that the utility of so doing exceeds the utility of waiting (here the year is construed as the age of the individual). The utility of retiring at age 65 is $V_{65} = U_D(D_{65})$, the only choice available at that time. The utility of working at age 64 is the value function $U_w\,(Y_{64}) + \beta U_D(D_{65})$, where b is a discount factor whose estimation is not always successful in these models; sometimes it is set outside the model (see additional note below). The utility of applying for disability benefits at age 64 is $U_D(D_{64})$ + b$U_D(D_{65})$. Note that D_{65} and all other incomes represented here depend on many factors related both to individuals and to the pattern of application, but we omit the list of conditioning variables to simplify exposition. All of the utilities involve disturbances. Finally, *Pr* [retire at 65 * no application before that] = 1.0 by assumption, while

$$Pr\,[\text{apply at 64} \mid \text{no application before that}\,] = P_r[(U_w\,(Y_{64}) + \beta\ U_D\,(D_{65}) \mid \text{retire at 65}) <$$

$$(U_D\ (D_{64}) + \beta\ U_D\ (D_{65}) \mid \text{apply at 64})].$$

That probability is a logit if the disturbances are distributed independent extreme values, and the random effect must be integrated out under the Butler and Moffitt (1982) method, as done by Daula and Moffitt (1995). Define V_{64} as the value of utility at age 64, including both possible paths, weighted by their probabilities.

We continue this backward one more period. The utility of working at age 63 is $U_w(Y_{63})$ + b$E_{63}V_{64}$. The utility of applying for disability benefits at age 63 is $U_D(D_{63}) + \beta U_D(D_{64}) + \beta^2 U_D(D_{65})$. Finally, Pr[apply at 63 * no application before age 63] = $Pr[(U_w(Y_{63})$ + bV_{64} * retire later) < ($U_D(D_{63})$ + b$U_D(D_{64})$ + b$^2U_D(D_{65})$ * retire at 63)]. The process continues with ever-growing formulae back to time t, thereby defining the probabilities for maximum likelihood or other methods of estimation each time V is defined. Eventually one returns to t, the present, and the first probability in time, but the last in calculation. In this last round we consider the probability that working in t has lower utility, including the value of choosing next period $(U_w\ (Y_t\ ) + \beta V_{t+1})$,

than applying for disability benefits now

$$(U_D(D_t\ ) + \beta U_D(D_{t+1}) + \ldots + \beta^{65-t}\ U_D(D_{65})).$$

The structural models contain the following parameters that we will estimate:

ß	is the discount rate or the rate of time preference. Beta equals one when there is no discounting.
k	is the value placed on a dollar of nonwage income relative to a dollar of wage income. It is less than one if stigma is associated with nonwage income.
ϒ	is a parameter that determines the marginal utility of money income. When gamma equals one the marginal utility of income is constant.
ρ	is a parameter determining the degree of autocorrelation of random utility in the LSW model.
σ_v^2 or σ_μ^2	is the variance of the persistent random component of utility.

σ_{ε}^{2} is the variance of the specific random component of utility. This value is assumed to be zero in the LSW option value model.

δ_w is a vector of coefficients used to predict wage income discussed below.

δ_n is a vector of coefficients used to predict nonwage income discussed below.