

Design of Efficient Social Insurance Institutions

John Rust *University of Maryland*

V.V. Chari *Federal Reserve Bank of Minneapolis*

Hugo Hopenhayn *UCLA*

prepared for conference on:

The Older Worker in the Labor Market: Work Time and Work Incentives

Lisbon, Portugal

May 15-17, 2005

Overview

1. Apologies to my discussant (Li Gan), paper late, NSF-funded work in progress.
2. Research focuses on two areas: 1) theory, and 2) computation
3. Computational focus aimed at yielding *practical* results, which would hopefully actually be useful to policymakers
4. Unfortunately, it is hopelessly naive to think that any theoretical or computational results will have any impact in the current political environment,
5. However the research *may* be useful when translated to firms, to help them design *optimal contracts* particularly to help deal with out of control escalation in health care costs.
6. Paper that has been circulated shows an initial, very limited foray into this area, where the “firms” in question are academic institutions.

Outline

1. Discussion of limits to theory and computation. However *parametric* approaches to mechanism design are feasible. Crawl before you walk.
2. Main practical problem with the latter approach is *credibility*: how do we know that rational decision making models are good models of how people actually behave?
3. Illustration: predicting the welfare and behavioral impacts of the Bush “individual accounts” reform plan.
4. Comments on the limited role of science in policy making and underinvestment in R&D in the US related to social insurance issues.
5. Illustration: “disability process reform” and the 1999 *Ticket to Work and Work Incentives Improvement Act*
6. Parametric approaches to optimal contract design in firms: results from “Impact of Retiree Health Plans on Faculty Retirement Decisions”

Using the Life Cycle Model to Evaluate the Impact of the “Bush Plan”

1. Q. What is the “Bush Plan”?
2. A. It is a “carve out” of 4.1 percentage points (1/3) of the 12.4% payroll tax, that individuals can elect to be placed in “personal retirement accounts”.
3. Bush Plan would initially cap contributions to PRAs at \$1000 but the cap would gradually increase over time.
4. Bush Plan does not allow early withdrawals from PRAs and mandates that upon retirement age, the amount in the PRA be drawn as a retirement annuity, together with traditional Social Security benefits.
5. *However the Bush adm has not been not specific about how traditional Social Security benefits would be reduced if someone chooses the PRA option.*

Impact of the “Bush Plan”, continued

No financing info provided in the White House’s February 2005 document, *Strengthening Social Security for the 21st Century*. The only information I could gather comes from a Google search.

“In February, the White House proposed an offset formula that would **reduce benefits by an amount tied to a worker’s account contributions plus a rate of return equal to that of the Treasury bond, currently estimated at 3 percent**. So a worker would need to earn more than 3 percent a year above inflation (assumed to be 3 percent) and expenses (estimated by the White House to be 0.3 percent) to come out ahead of just staying in the traditional system. That means his investment account would need to earn more than 6.3 percent, not adjusted for inflation. That’s not an impossible hurdle, but it’s also not an easy one.”

Impact of the “Bush Plan”, continued

“Administration officials may consider proposing a lower offset rate, according to an interview in the Wall Street Journal with Allan Hubbard, the director of the National Economic Council and assistant to the president for economic policy.”

quoted from CNN Money online article “Bush to change account formula?” March 24, 2005 http://money.cnn.com/2005/03/24/retirement/offset_change

- In view of these ambiguities, I am going to take the liberty of filling in the details in order to simulate the effects of my own take on the “Bush plan”
- My version of the Bush Plan is more **libertarian**. In my version workers can withdraw money from their PRAs at any time, and in particular, *they are not forced to annuitize their PRA balances*. However my calculations are also consistent with **compassionate conservatism** that is, *I assume that disability benefits remain untouched by the reform*.

Impact of the “Bush Plan”, continued

- With *generational accounting* in mind, I assume that *cohort neutral financing* is adopted. That is, I reduce benefits so that in present value terms, there is no net transfer or outflow to the cohort as a result of the carve out.
- I assume there will be a proportionate cut in benefits with a benefit reduction factor set so the expected present value of benefits (discounted back to age 20 at a 3% discount rate) equals the reduction in the present discounted value of Social Security contributions.
- I make my calculations under the hypothesis that the Bush plan had been enacted, (retroactively), for the 1930-1940 birth cohort, since I have data on complete earnings histories and other key information, courtesy of the **Health and Retirement Survey**.
- My calculations indicate that in order to enact cohort neutral financing of the Bush plan, *the 33% reduction in Social Security contributions must be matched by a 50% reduction in Social Security Old Age benefits.*

Impact of the “Bush Plan”, continued

- Q. Why is the reduction in benefits a larger percentage reduction than the 1/3 reduction in contributions?
- A. Because benefit payments are far in the future, whereas contributions are made “up front”.
- This implies that the present value of benefits, being further away in the future, and taking mortality into consideration, is smaller in present value than the up front contribution stream.
- Thus, if the contribution stream is reduced by 1/3, the required reduction in the benefit stream must be larger than 1/3 *if the rate of return used to discount future cash flows is larger than the internal rate of return on Social Security for the cohort in question.*

- For the 1930-1940 birth cohort, the “aggregate” IRR (i.e. the interest rate that makes the present value of all contributions equal to the present value of all benefits, including DI benefits) is 1.2%. Thus, if the government was using a 1.2% discount rate, then a 1/3 reduction in contributions could be balanced by a 1/3 reduction in benefits.
- However if the government can actually “invest” at a higher long term rate of return, e.g. 3%, this is the appropriate rate to discount each cohort’s contributions and benefits. *My calculations indicate that if the Bush plan were to reduce benefits by only 1/3 along with a 1/3 reduction in contribution, this would be equivalent to a per capita net transfer of \$8000 to each individual in my simulation cohort.*
- My simulations of a life cycle model are based on a fairly realistic model of the Social Security benefit formula. In particular the model keeps track of AIME and uses the appropriate cohort-specific benefit rules to compute the PIA’s, actuarial benefit reduction factors, and “earnings test” that were applicable to the 1931-1941 birth cohort.

- On an undiscounted basis, my simulations indicate that a typical member of this cohort contributes an average of \$119,800 in Social Security taxes (i.e. the 12.4% employee and employer combined contributions) over their lifetime.
- On an undiscounted basis, my simulations indicate that a typical member of this cohort receives an average of \$192,800 in Social Security benefits over their lifetime, including Social Security disability benefits.
- However on a discounted basis, (3% discount rate), the present value of contributions per capita is \$65,853 and the present value of benefits is \$41,378, *members in this cohort are actually providing a net subsidy of \$24,475 to the Social Security system on a present discounted value basis.*
- Under my simulation of the Bush plan, with a 50% reduction in Social Security benefits, the per capita expected present value of contributions is \$45,420 and the expected present value of benefits is \$22,180, *so the 50% reduction in benefits has succeeded in approximately equalizing the net transfers that this cohort makes to the rest of the Social Security System.*

Using a Calibrated Life Cycle Model to Predict the Impact of Bush Plan

- Model starts at age 20, in annual periods to maximum age of 100
- Individuals choose 1) consumption, 2) labor supply, and 3) Social Security application decision, including disability
- Individuals are subject to health shocks, uncertain mortality, and earnings shocks. We use actual Social Security records to produce “realistic” model of earnings dynamics.
- We provide a detailed treatment of the Social Security rules including AIME, PIA, DRC, earnings tests, DI application/appeal procedure, continuing disability reviews, trial work period, and Extended Period of Eligibility.
- Individuals have a bequest motive, and utility that is additively separable in consumption and leisure, with CRRA preferences for consumption

$$u(c, l) = [c^\gamma - 1]/\gamma + \log(l) \quad (1)$$

with $\gamma = -.37$ (moderate risk aversion).

Using a Calibrated Life Cycle Model to Predict the Impact of Bush Plan

- Individuals have discount factor $\beta = .92$, and can invest in (riskless) assets for a pre-tax return of 5% and an after-tax return of 3.75%.
- Mortality probabilities and earnings dynamics derived for individuals in the 1930-1940 birth cohort from the Health and Retirement Survey, with other parameters of preferences and technology (health-specific award and audit rates, and stigma effects and “hassle costs” associated with being on or applying for DI) to match observations on entry and exit to DI roles.
- I solved life cycle model under a *status quo* scenario where the members of the cohort received the benefits under the actual Social Security rules in effect for them, and under the hypothesis that at age 20 they were offered the “Bush plan”.

Simulation Results

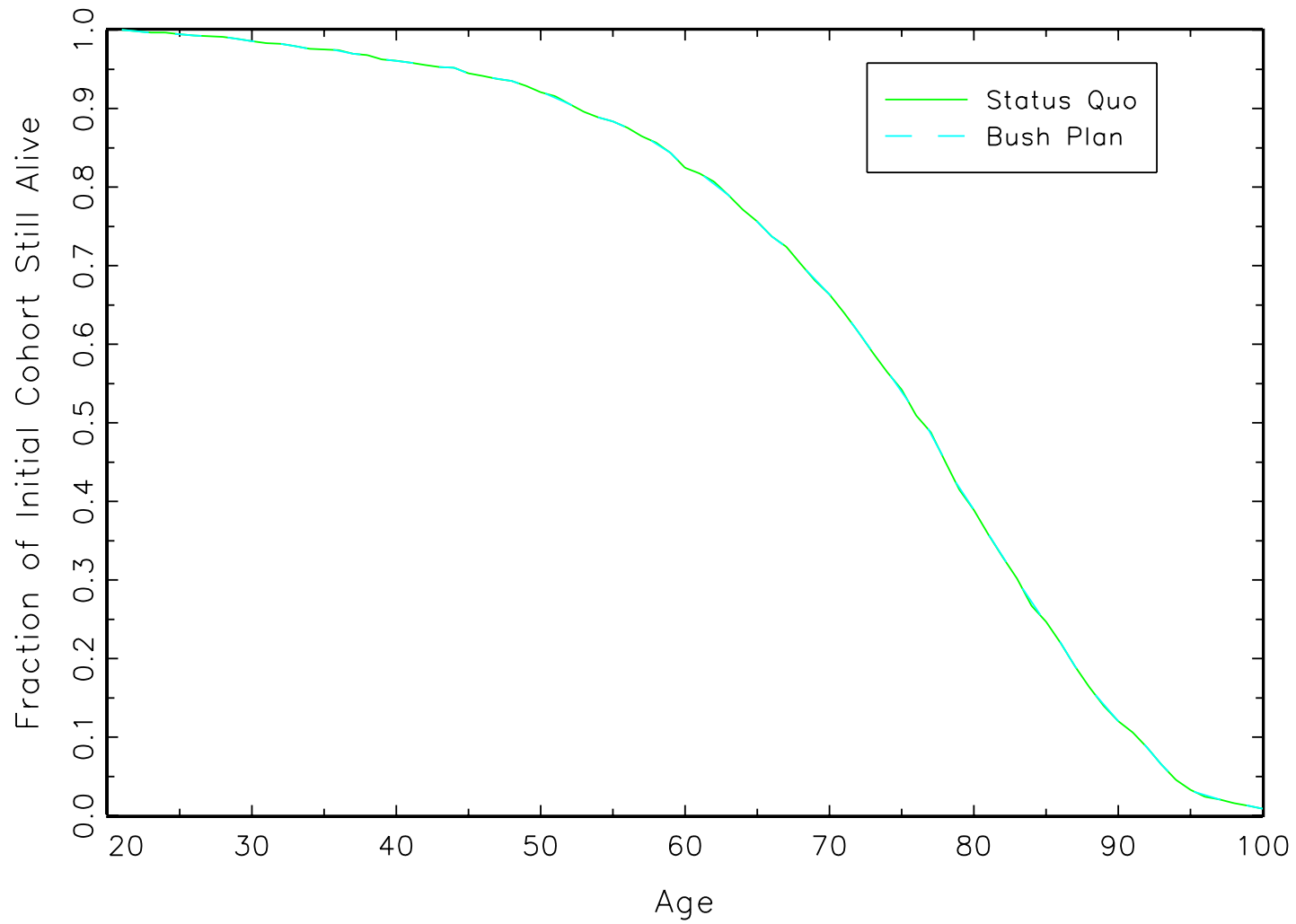
- the DP model predicts that at all ages up to 50, all individuals are better off (in expected discounted utility terms) under the Bush plan compared to the *status quo*.
- using the calculated optimal decision rules, I simulated life histories for 1123 starting from random initial conditions (initial wealth and health) at age 20, followed until they die under the *status quo* and Bush plan scenarios.
- individuals work more, retire later, accumulate more wealth, leave more bequests, and consume more under Bush plan than *status quo* except for individuals who are unlucky enough to live beyond 90. Those old folks are worse off under Bush plan.
- In general there is *ex post envy* under Bush plan: after age 50, individuals would prefer to switch back to *status quo* if it could be done at no cost to them. There is also greater inequality, and higher incidence of old age poverty under the Bush plan.

“Bush Plan” versus Status Quo for 1931-1941 Cohort

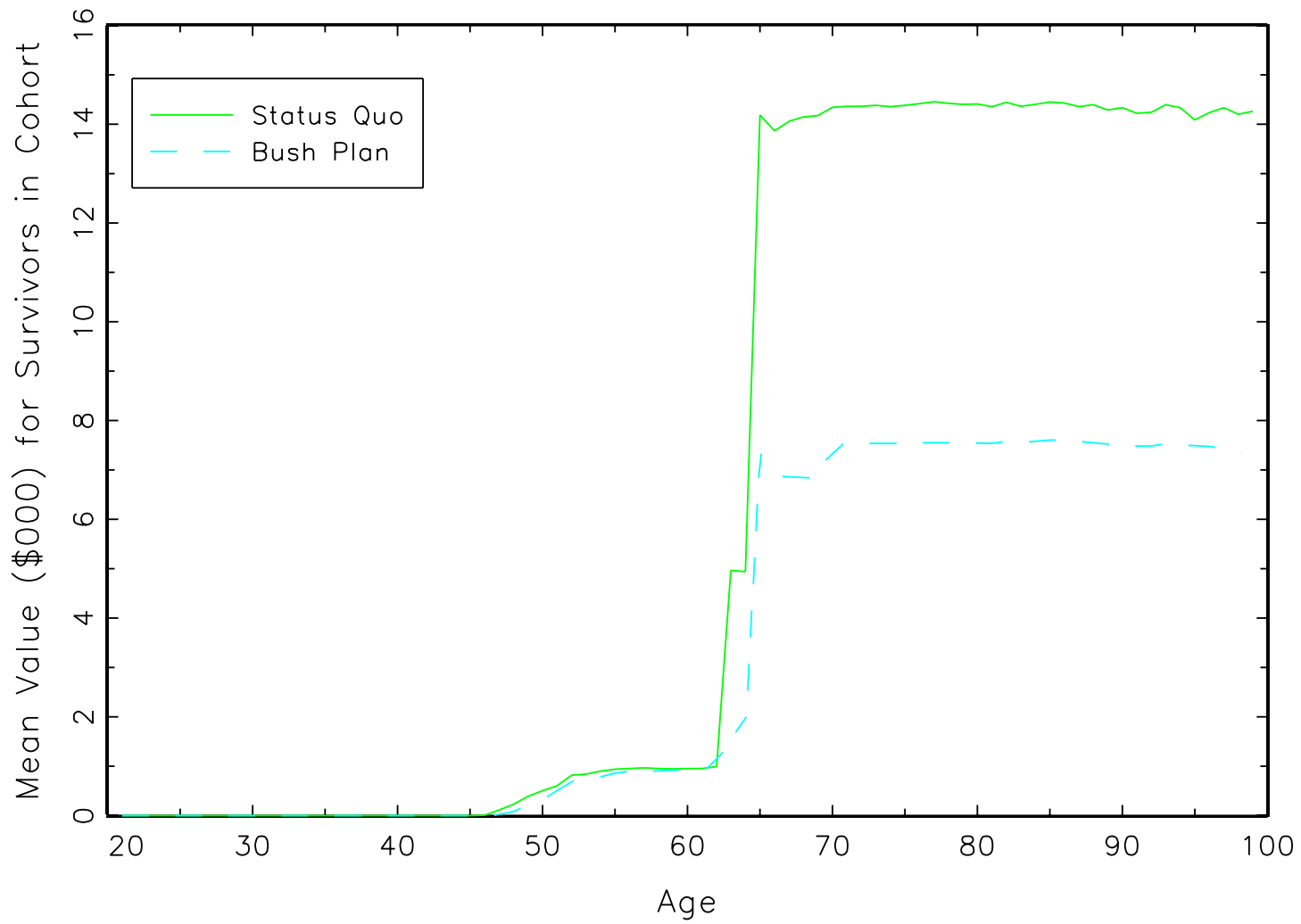
(Dollar amounts are per capita present values (3%), in thousands)

Item	Bush Plan	Status Quo
Federal Tax payments	99.09	93.84
Social Security contributions	45.42	65.85
Social Security benefits	22.18	41.4
Net transfer from this cohort	23.24	24.45
Wage earnings	550.75	532.39
Consumption	468.03	450.07
Fraction of Sample who Apply for DI	10.7%	11.9%
Fraction on DI	8.1%	8.5%
Total Person years on DI	1113	1264
Fraction of DI recipients who return to work	46.1%	9.5%

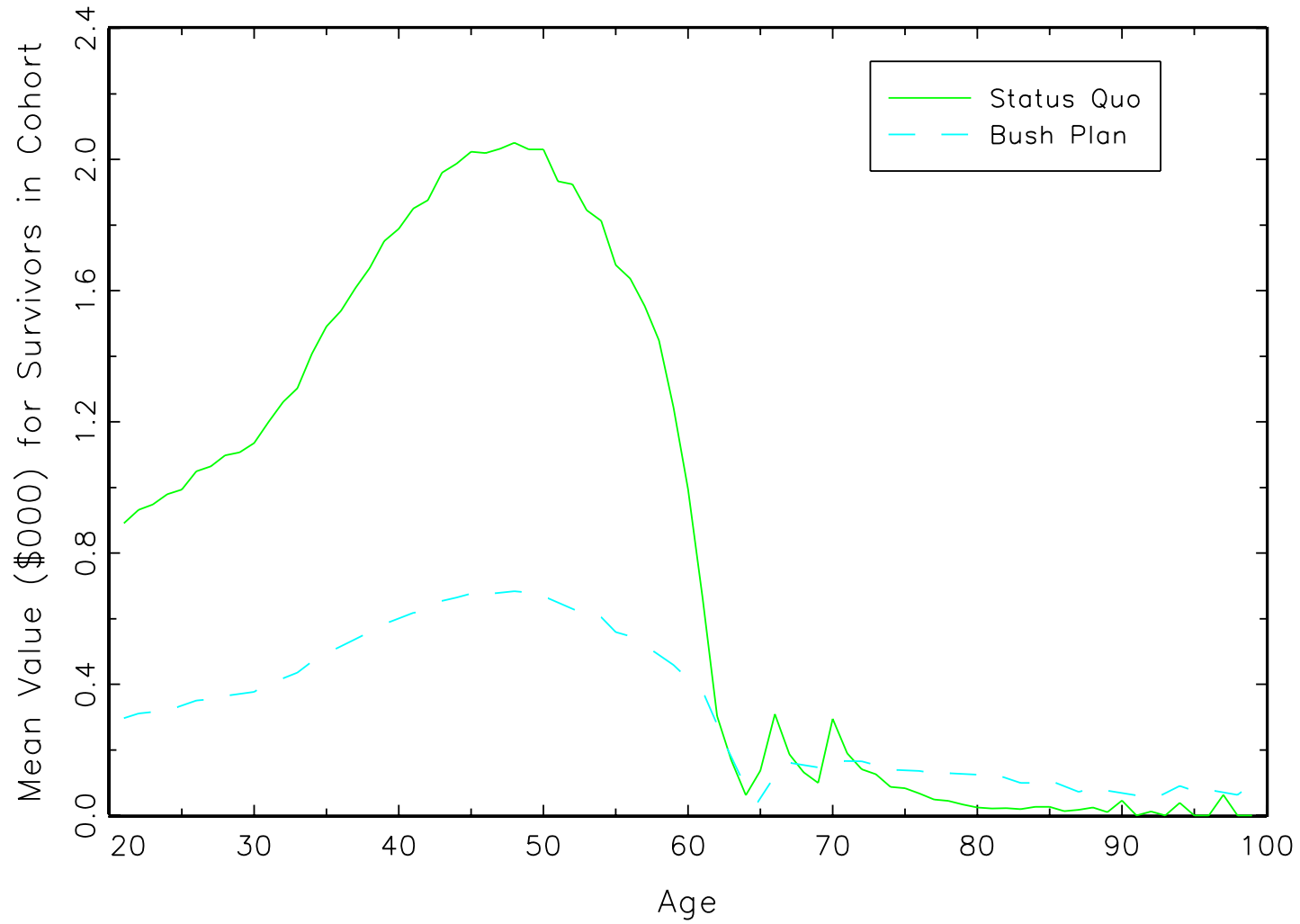
Simulated Impact on Survival Rates



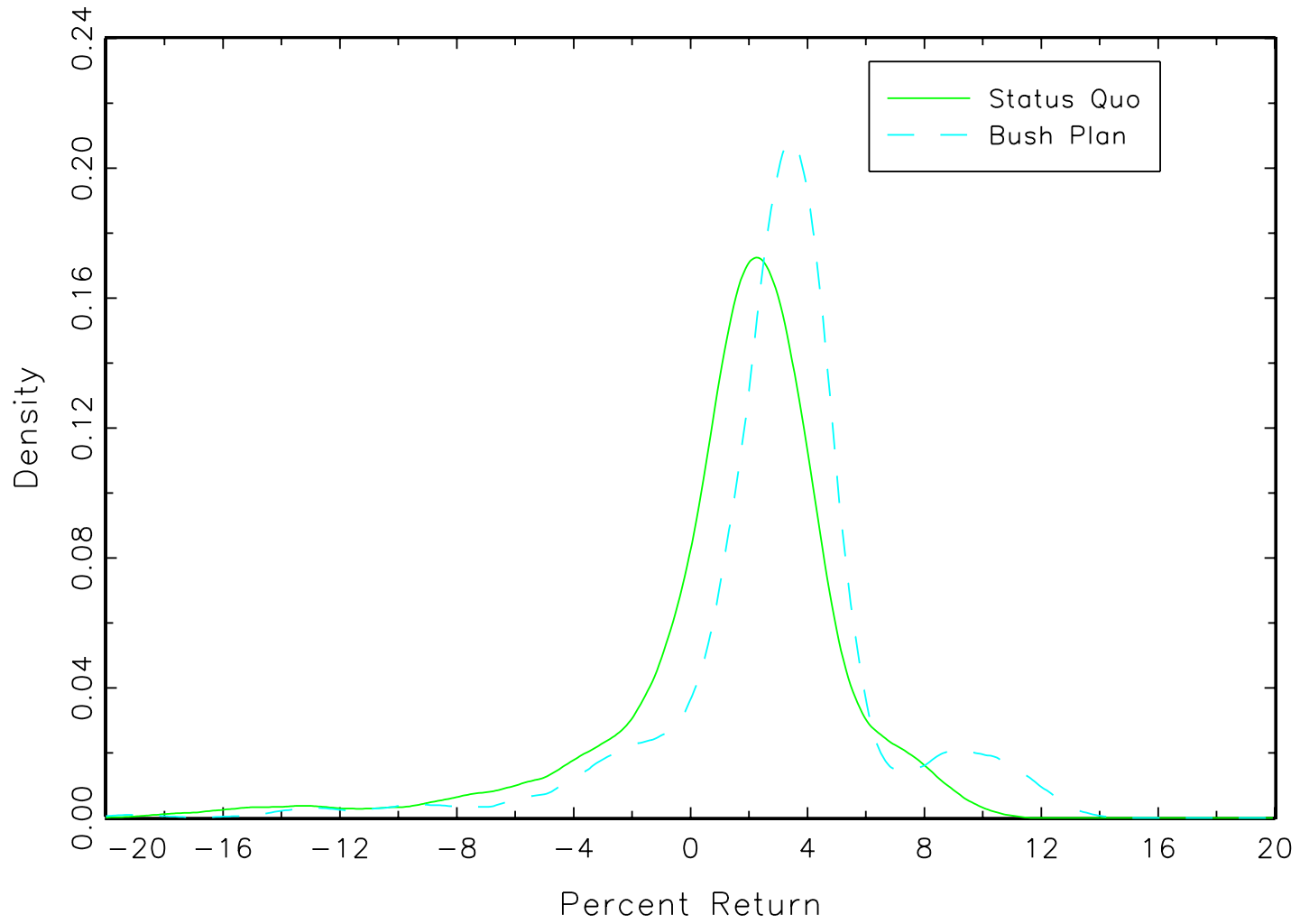
Impact on Social Security Benefits



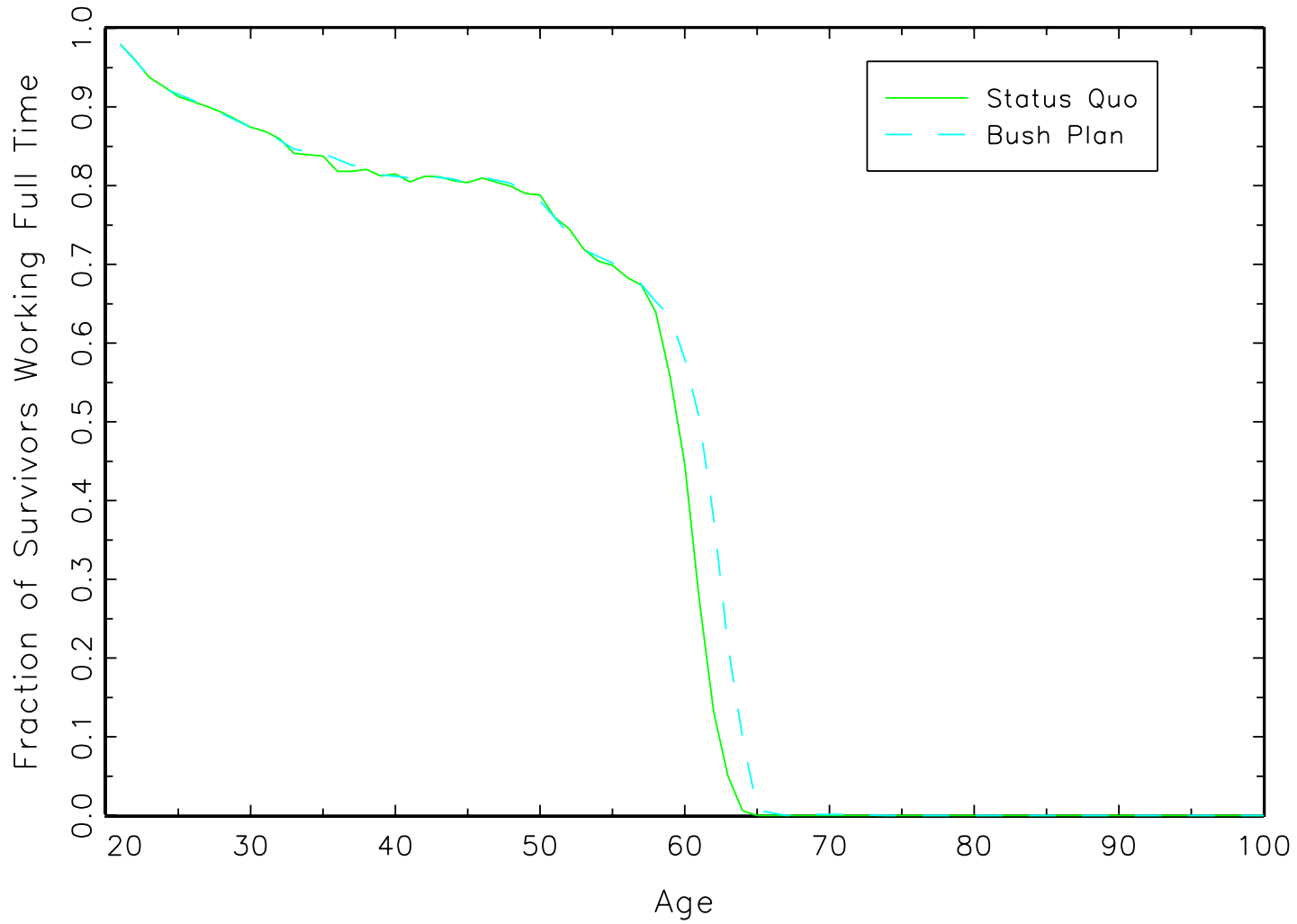
Impact on Social Security Contributions



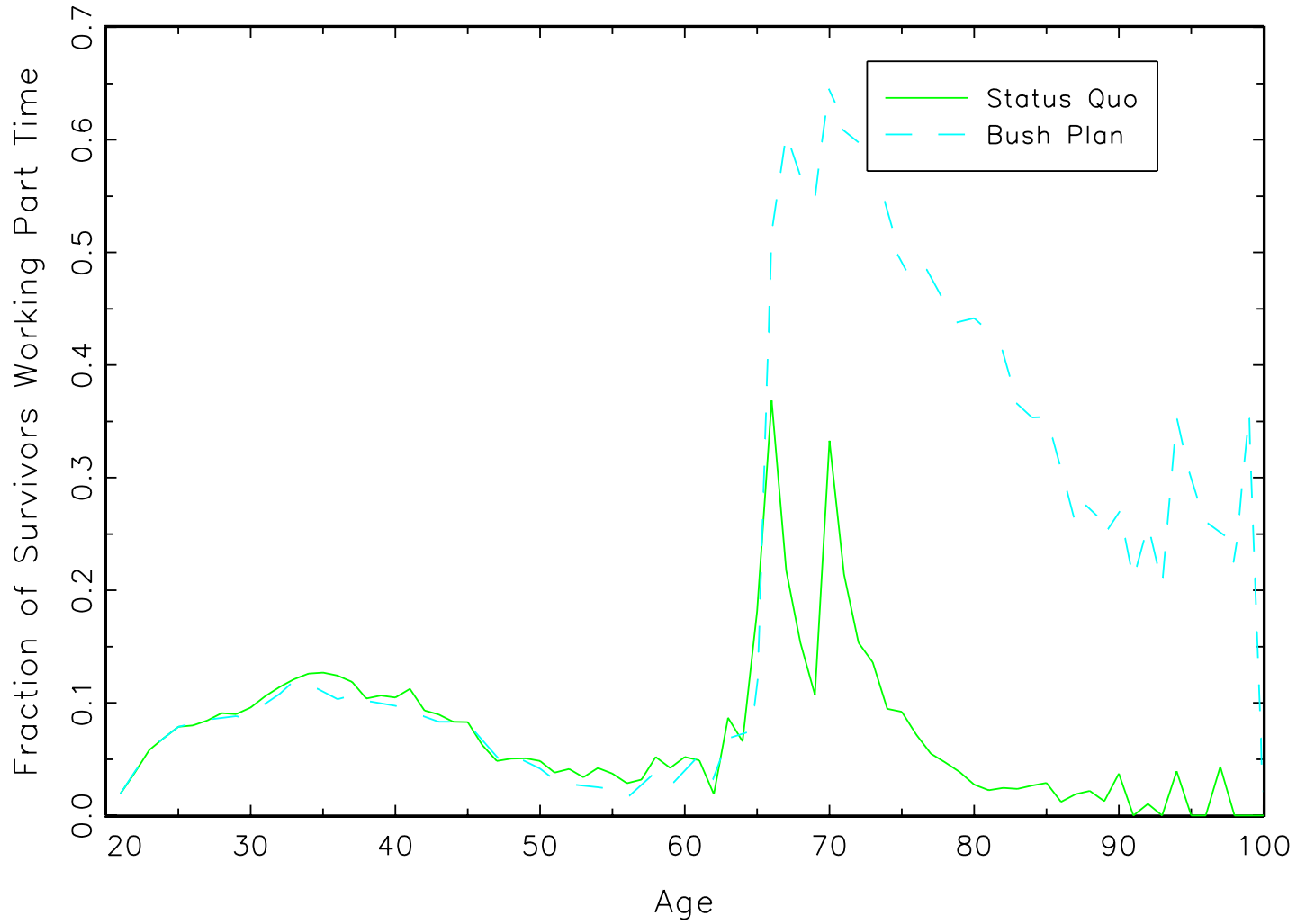
Impact on Distribution of IRRs on SS Contributions



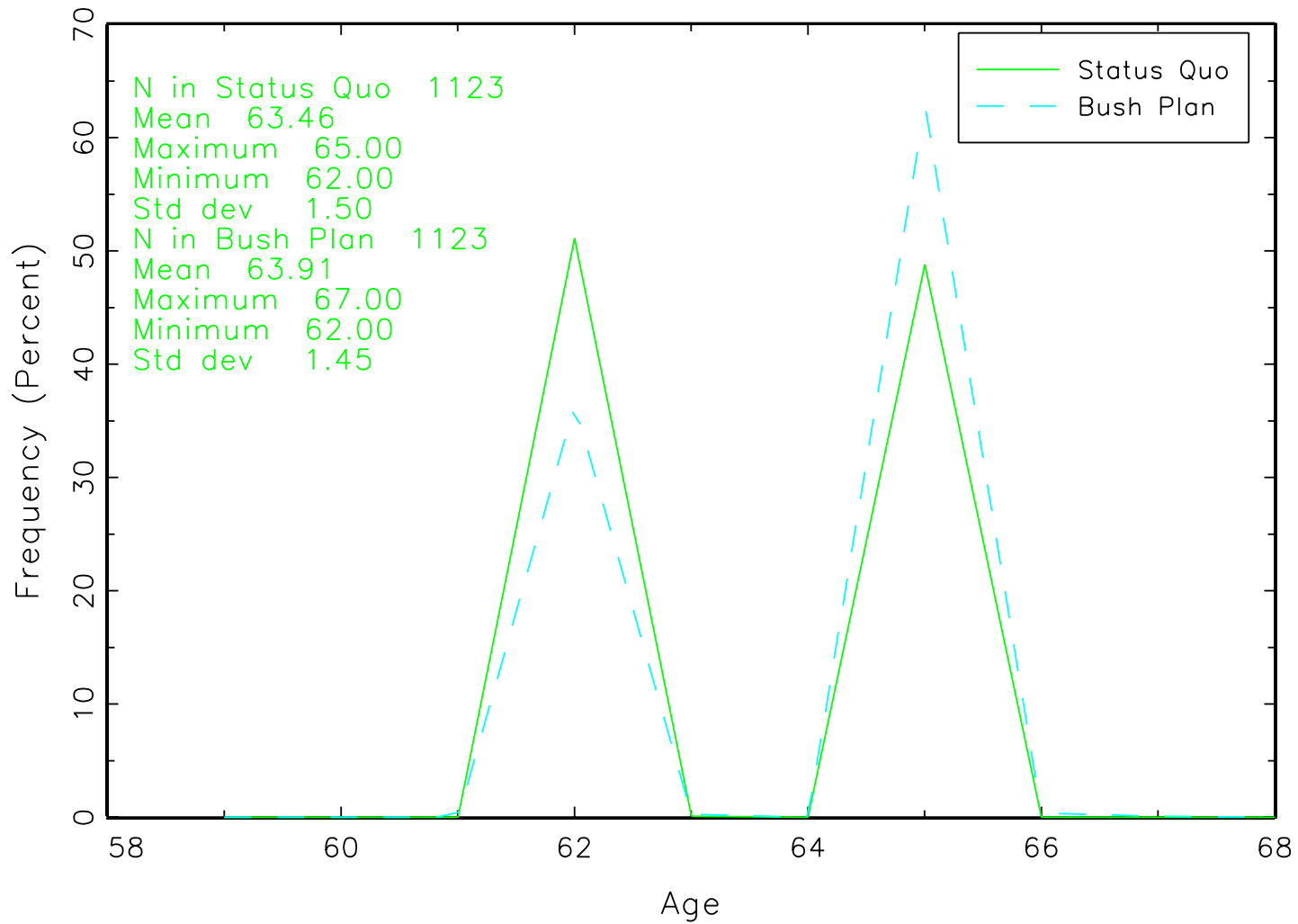
Impact on Full Time Employment



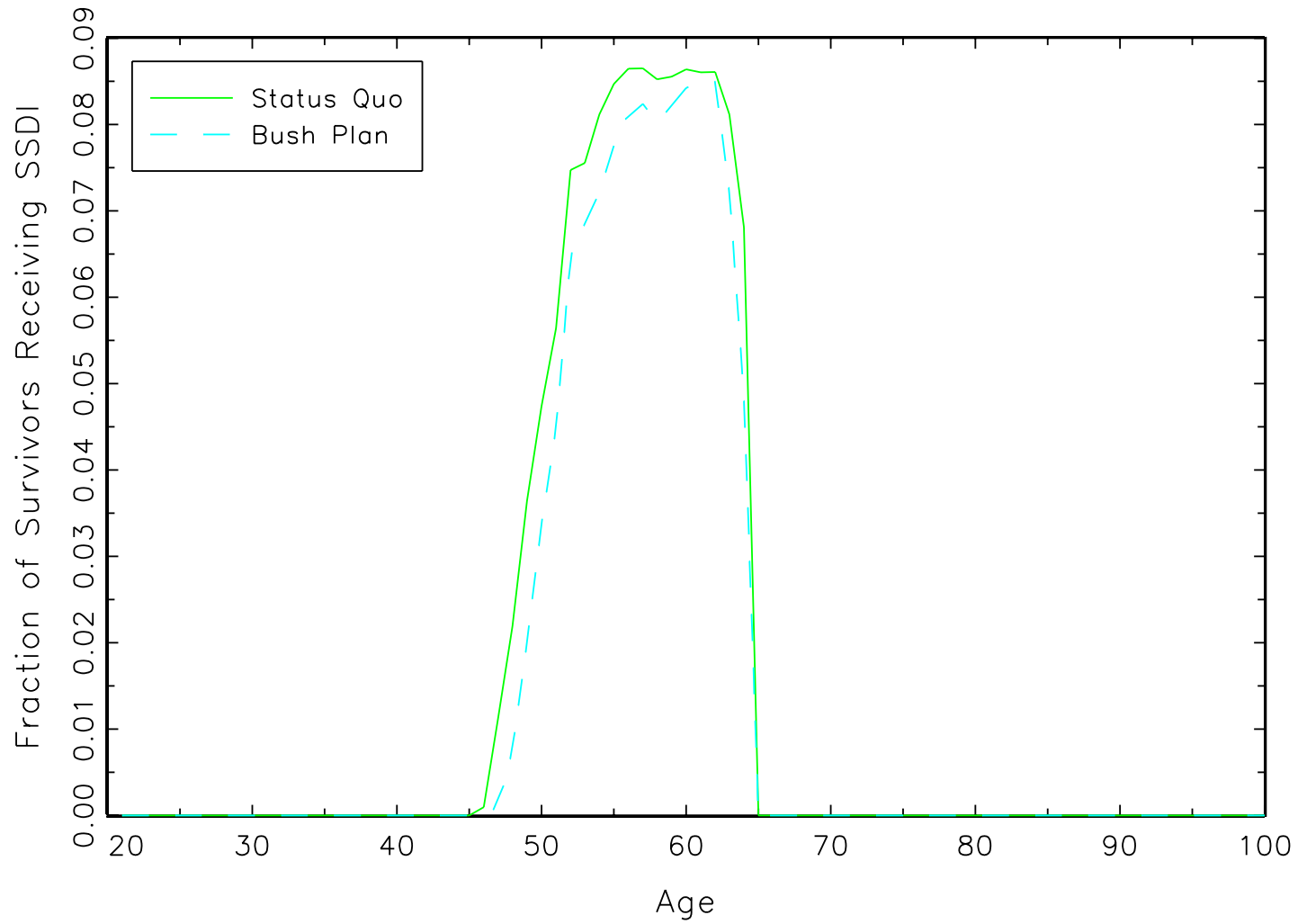
Impact on Part Time Employment



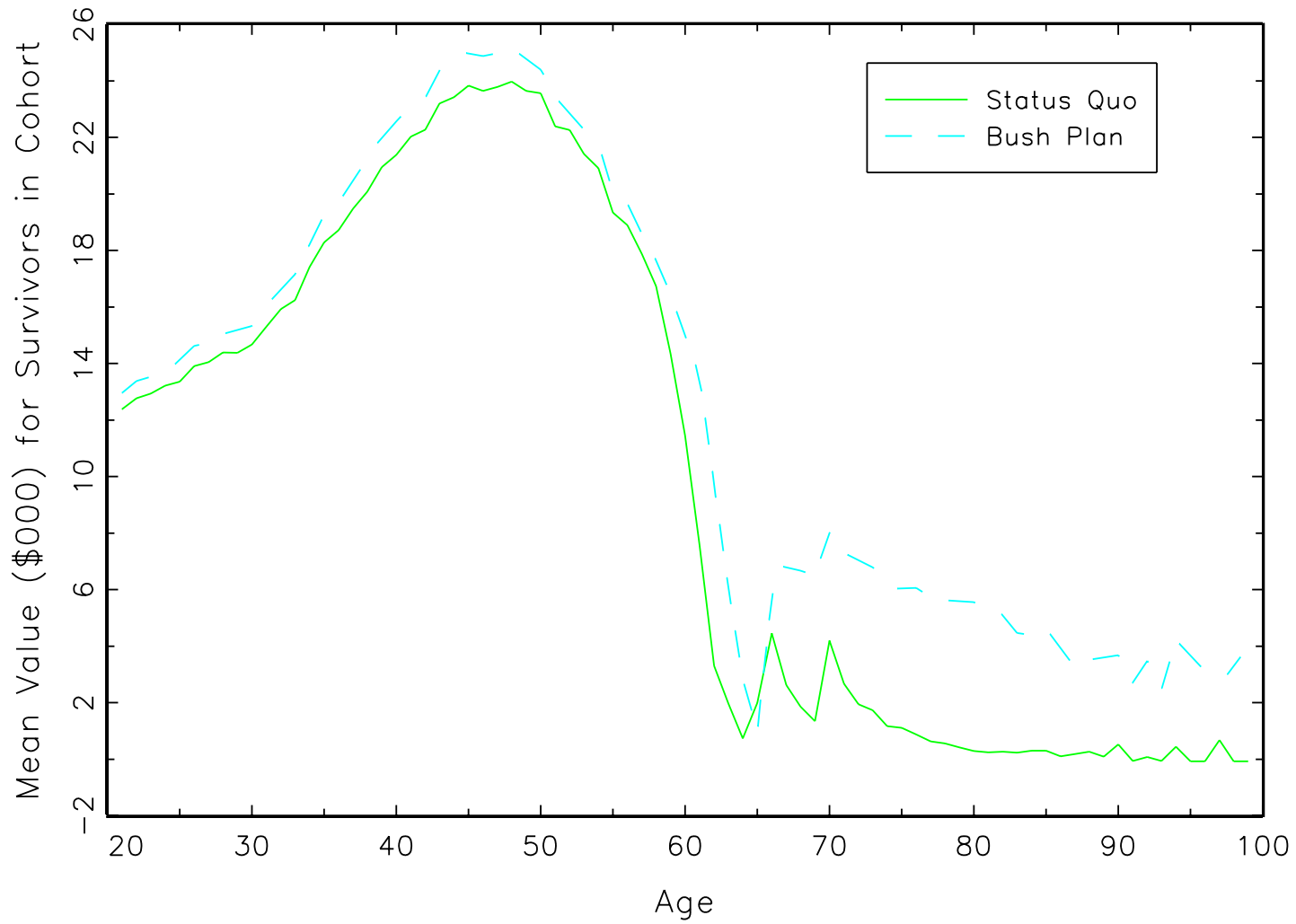
Impact on Ages of First Entitlement to OA Benefits



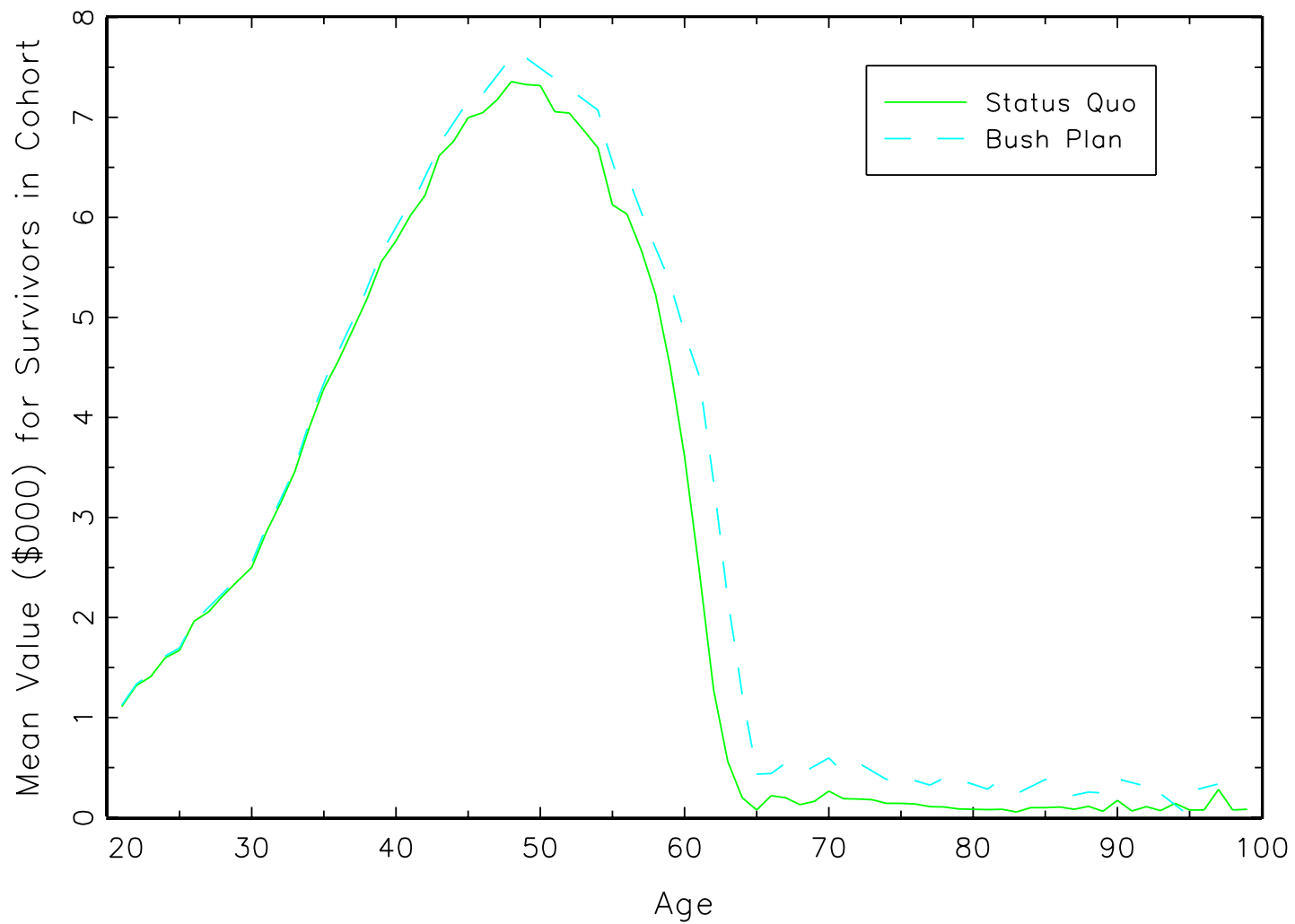
Impact on Disability Roles



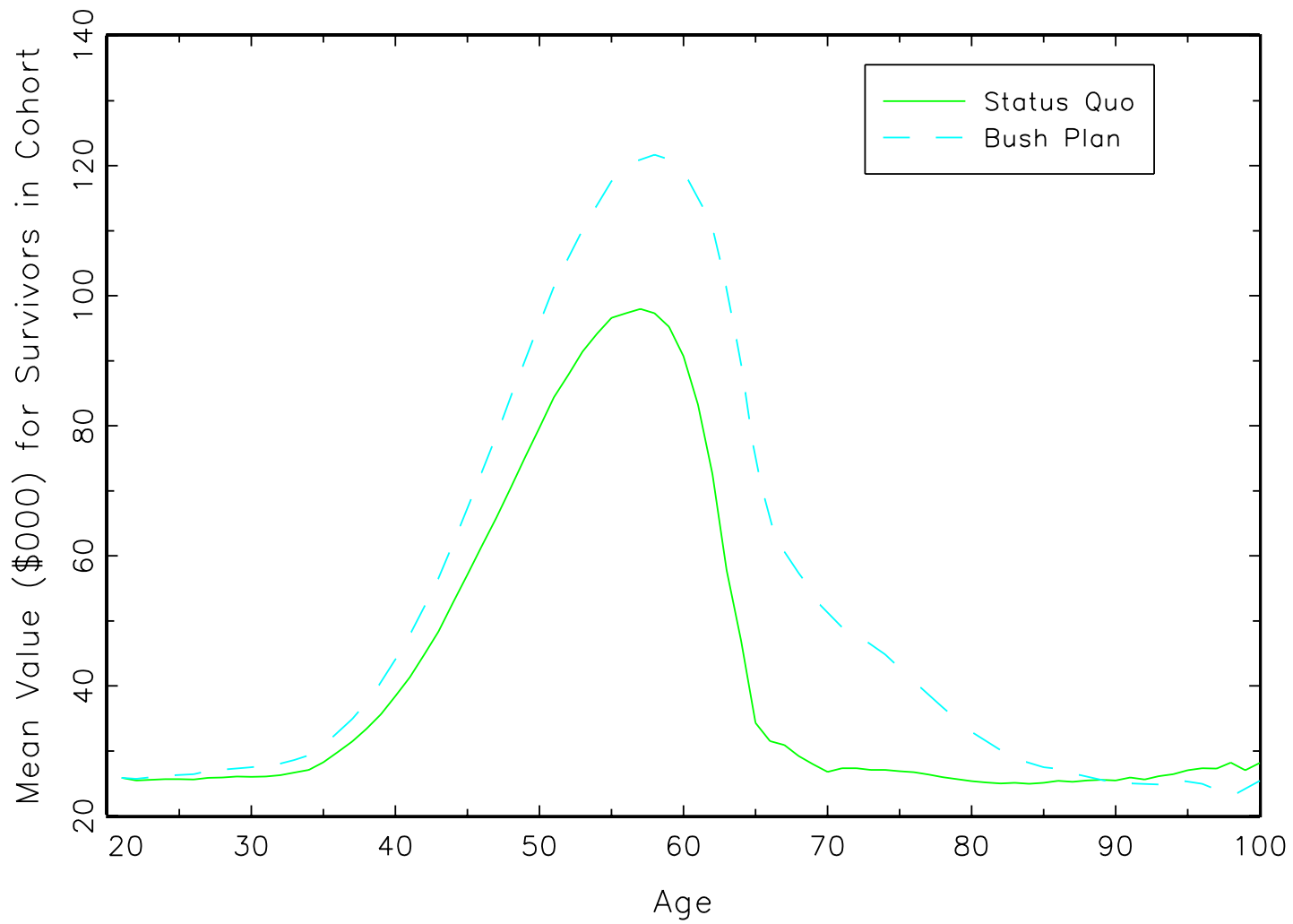
Impact on After Tax Wages



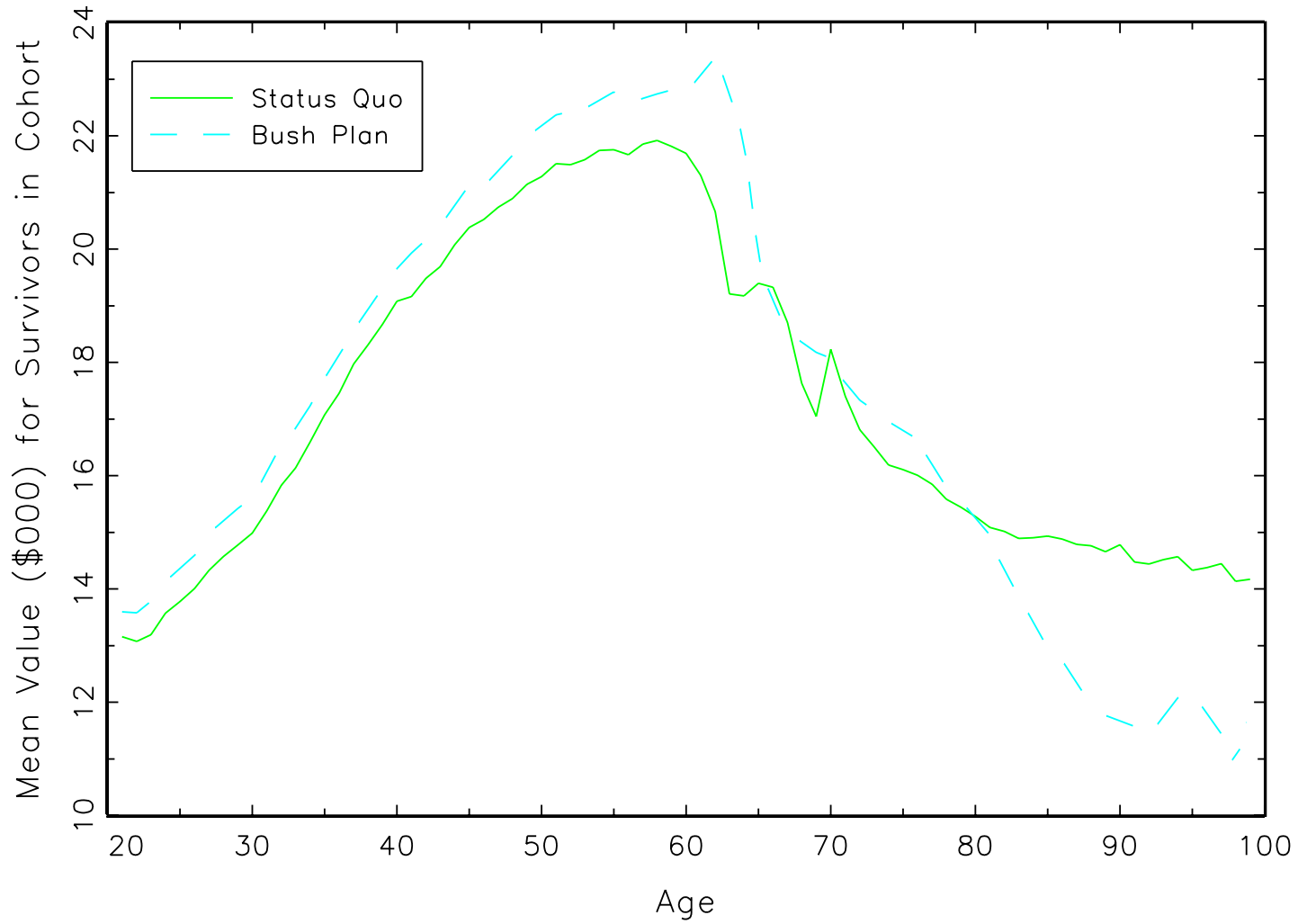
Impact on Federal Tax Payments



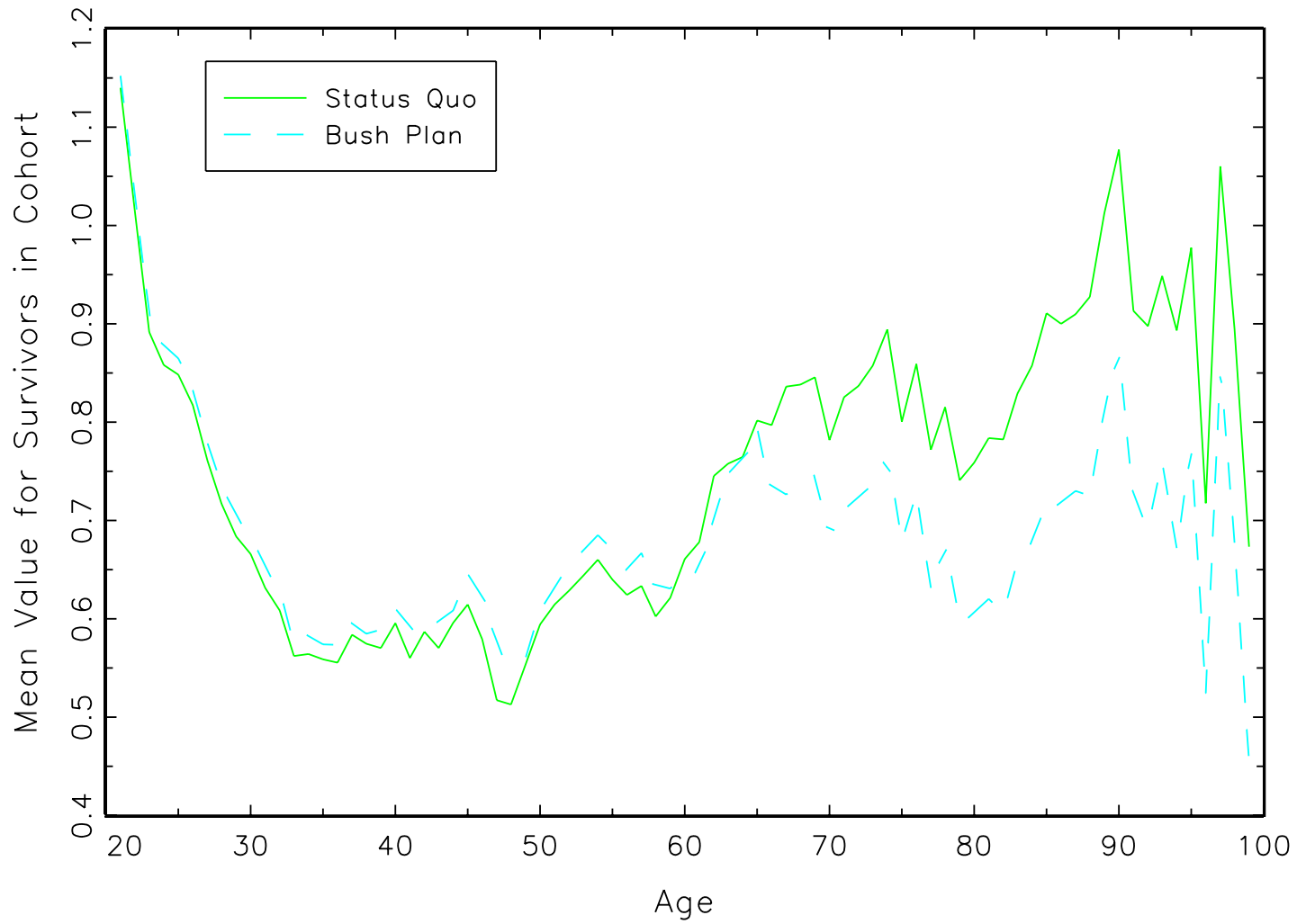
Impact on Wealth Accumulation



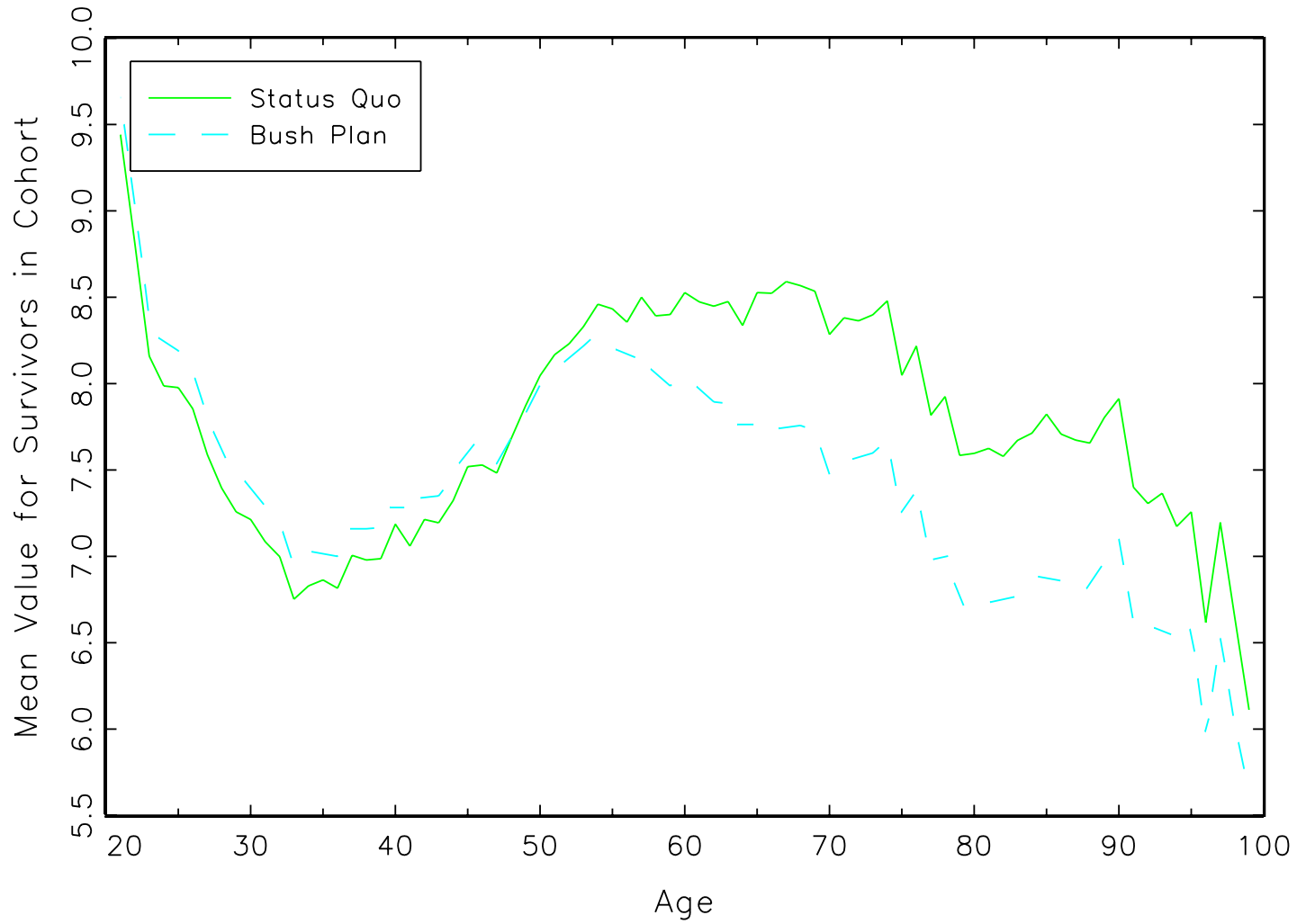
Impact on Consumption



Impact on Current Period Utility



Impact on Expected Discounted Utility



**How Large are the Classification Errors in the Social Security Disability
Award Process?**

John Rust, *University of Maryland*

Hugo Benítez-Silva, *SUNY-Stony Brook*

Moshe Buchinsky, *UCLA, NBER, and CREST-INSEE*

Disability at a Glance

- **10 million beneficiaries** SSDI and SSI, 3.3% of U.S. population
- **Total cost, \$106 bn/year** (excluding Medicare/Medicaid), nearly 1% of GDP
- **High Administrative costs:** responsible for 90%, at a cost of \$2.25 billion/year in 54 Disability Determination Service Centers (DDS), with 15,000 employees processing about 4 million applications per year
- **Unsustainable growth** Since 2001:
 - applications have been growing at 13% per year,
 - awards have been growing at 7% per year,
 - roles have been growing at 5% per year.
- *If the rapid growth in DI roles persists (5 times faster than the US population as a whole), along with the alarming growth in obesity, by 2075 over 90% of the U.S. population will be obese and disabled!*

How Large are the Classification Errors?

- Our point estimate that the fraction of **award errors**, i.e. the fraction of SSDI/SSI applicants who are ultimately awarded benefits and who are *not* “disabled” is **28%**
- Our point estimate that the fraction of **rejection errors**, i.e. the fraction of SSDI/SSI applicants who are rejected and who *are* “disabled” is **61%**

How Do We Get These Estimates?

- We use the *Health and Retirement Survey* (HRS) and follow a sample of 12,000+ older Americans between 1992 and 1998 (first 4 waves of the HRS).
- We compare their *self-reported disability status* \tilde{d} to the SSA's *ultimate award decision* \tilde{a}
- We argue that individuals are truthful, accurate reporters of their “true” disability status $\tilde{\tau}$, so by comparing \tilde{a} and \tilde{d} we can infer error rates in the SSA's bureaucratic award and appeal process.

Two Similar Definitions of “Disability”

- **SSA** *The inability to engage in any substantial gainful activity (SGA) by reason of any medically determinable physical or mental impairment, which can be expected to result in death, or which has lasted, or can be expected to last, for a continuous period of at least 12 months.*
- **HRS** *Do you have a health condition that prevents you from working entirely?*

Econometric Analysis of Classification Errors

- Define three binary random variables \tilde{a} , \tilde{d} and $\tilde{\tau}$ as follows:

$$\tilde{a} = \begin{cases} 1 & \text{if person is ultimately awarded SSDI/SSI benefits} \\ 0 & \text{otherwise} \end{cases}$$

$$\tilde{d} = \begin{cases} 1 & \text{if a health problem prevents one from working entirely} \\ 0 & \text{otherwise} \end{cases}$$

$$\tilde{\tau} = \begin{cases} 1 & \text{if someone is "truly disabled" according to the SSA definition} \\ 0 & \text{otherwise} \end{cases}$$

Definitions of Classification Errors

- **Award Error Rate** This is the probability a person is *not* truly disabled given that they are awarded benefits, $\Pr\{\tilde{\tau} = 0|\tilde{a} = 1\}$.
- **Rejection Error Rate** This is the probability a person *is* truly disabled given that they were rejected, $\Pr\{\tilde{\tau} = 1|\tilde{a} = 0\}$.
- Note that the award and error rates differ from, but are related to, the Type I and II error rates in hypothesis testing.
- **Type I error rate** the probability a person is rejected given that they are truly disabled, $\Pr\{\tilde{a} = 0|\tilde{\tau} = 1\}$
- **Type II error rate** the probability a person is accepted given that they are not truly disabled, $\Pr\{\tilde{a} = 1|\tilde{\tau} = 0\}$.

Consider the “Easy” Case First

- Suppose that $\tilde{\tau} = \tilde{d}$ with probability 1, i.e. that individuals know and truthfully report their truly disability status, using the current “social standard” of disability for the current socio/political environment.
- Then our point estimates for SSA’s error rates are as follow
- **Award Error Rate** $\Pr\{\tilde{d} = 0|\tilde{a} = 1\} = .28$
- **Rejection Error Rate** $\Pr\{\tilde{d} = 1|\tilde{a} = 0\} = .61$
- **Type I Error Rate** $\Pr\{\tilde{a} = 0|\tilde{d} = 1\} = .26$
- **Type II Error Rate** $\Pr\{\tilde{a} = 1|\tilde{d} = 0\} = .63$

Now Consider the “Harder” but More Realistic Case

- Assume that both \tilde{a} and \tilde{d} are noisy but unbiased indicators of true disability status $\tilde{\tau}$
- Assume that we can model $(\tilde{a}, \tilde{d}, \tilde{\tau})$ as a trivariate probit with a correlation structure designed to match the correlation between the observed random variables \tilde{a} and \tilde{d} .
- Under these assumptions we can estimate the parameters of the trivariate probit model by maximum likelihood and use the resulting model to infer the classification and Type I and II error rates using **Bayes Rule**.
- Surprisingly, when we do these computations in this more realistic case, the rate of classification errors and the Type I and II error rates differ by only a small amount from the error rates we obtained in the “easy” case when we assumed that $\tilde{d} = \tilde{\tau}$.

Bayes Estimates of the Classification Errors

- **Award Error Rate:** $\Pr\{\tilde{d} = 0 | \tilde{a} = 1\} = .23$
- **Rejection Error Rate** $\Pr\{\tilde{d} = 1 | \tilde{a} = 0\} = .61$
- **Type I Error Rate** $\Pr\{\tilde{a} = 0 | \tilde{d} = 1\} = .23$
- **Type II Error Rate** $\Pr\{\tilde{a} = 1 | \tilde{d} = 0\} = .68$

Previous “Audits” of SSDI Award Process Yielded Similar Conclusions
using very different methodologies

Summary of DI Classification Errors from Nagi (1969) Nagi compared an “expert decision” (a moderated group decision of an examining team consisting of a physician, psychologist, social worker, occupational therapist, and a vocational rehabilitation expert) to SSA’s award decision

Expert Team Decision	SSA Award Decision		Total
	Awarded	Denied	
Can Work	291 (19.3%)	492 (52.1%)	783 (31.9%)
Cannot Work	1,219 (80.7%)	452 (47.9%)	1,671 (68.1%)
Total	1,510 (61.5%)	944 (38.5%)	2,454 (100.0%)

Our version of Nagi (based on a subsample of 360 HRS respondents for which complete information on ADLs and health characteristics are available)

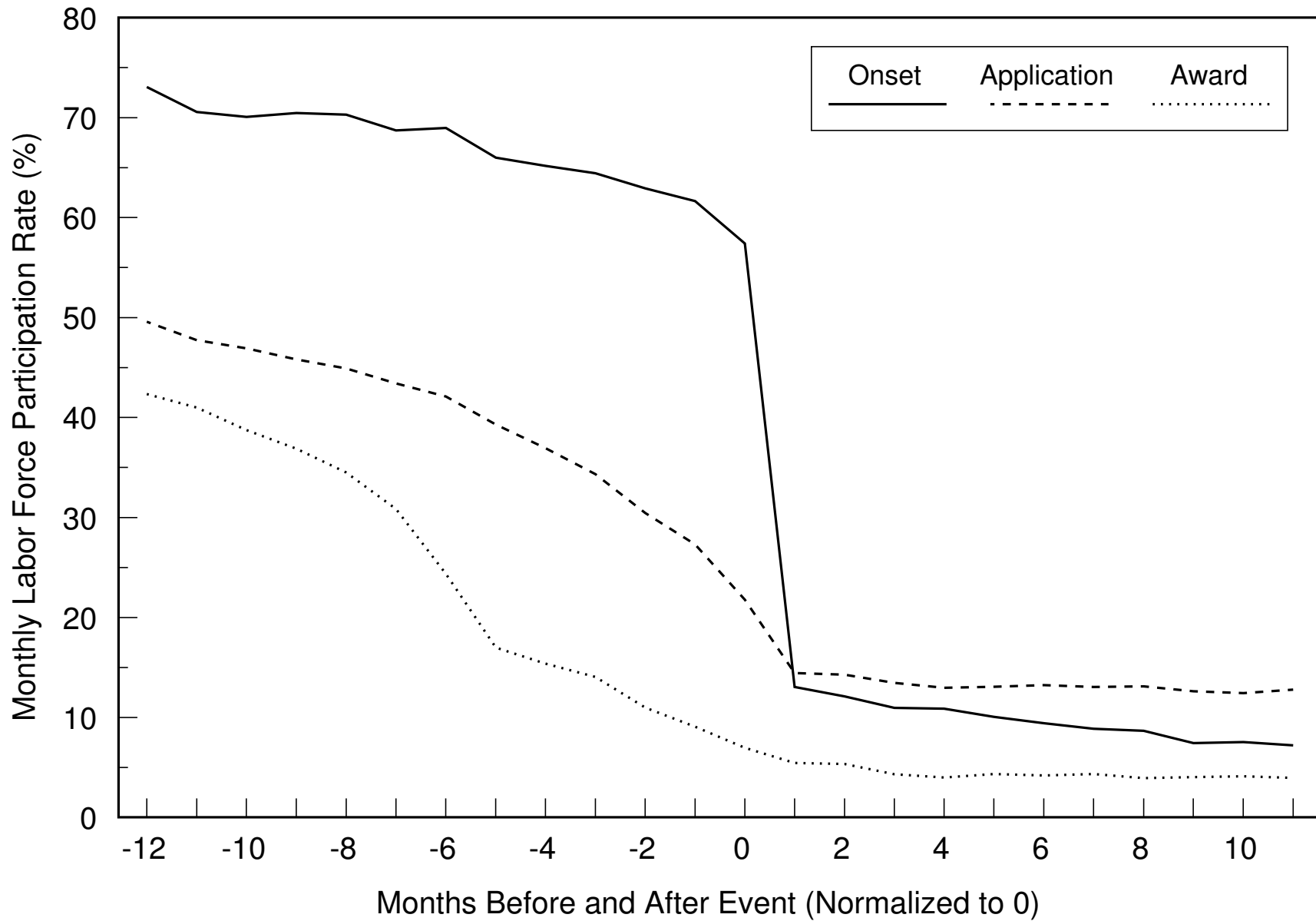
We don't have access to an independent expert, so we compare self-reported disability to SSA's award decision

Self-Reported Disability Status	SSA Award Decision		Total
	Awarded	Denied	
Not Disabled	60 (22.6%)	42 (44.7%)	102 (28.3%)
Disabled	206 (77.4%)	52 (55.3%)	258 (71.7%)
Total	266 (73.9%)	94 (26.1%)	360 (100.0%)

GAO study critical of high reversal rates in appeals of initial denials to ALJs

Are the inconsistencies between first stage rejections by DDSs and subsequent awards by ALJs symptoms of decision errors? If so, are the DDSs too strict or are the ALJs too lenient?

Condition	DDS Award Rate	ALJ Award Rate
Physical	29%	74%
Musculoskeletal	16	75
Back cases	11	75
Other	23	76
Other physical	36	74
Mental	42	87
Illness	39	87
Retardation	54	84
All impairments	30	77



Testing Unbiasedness of Self-Reported Disability

- **Unbiasedness Test:**

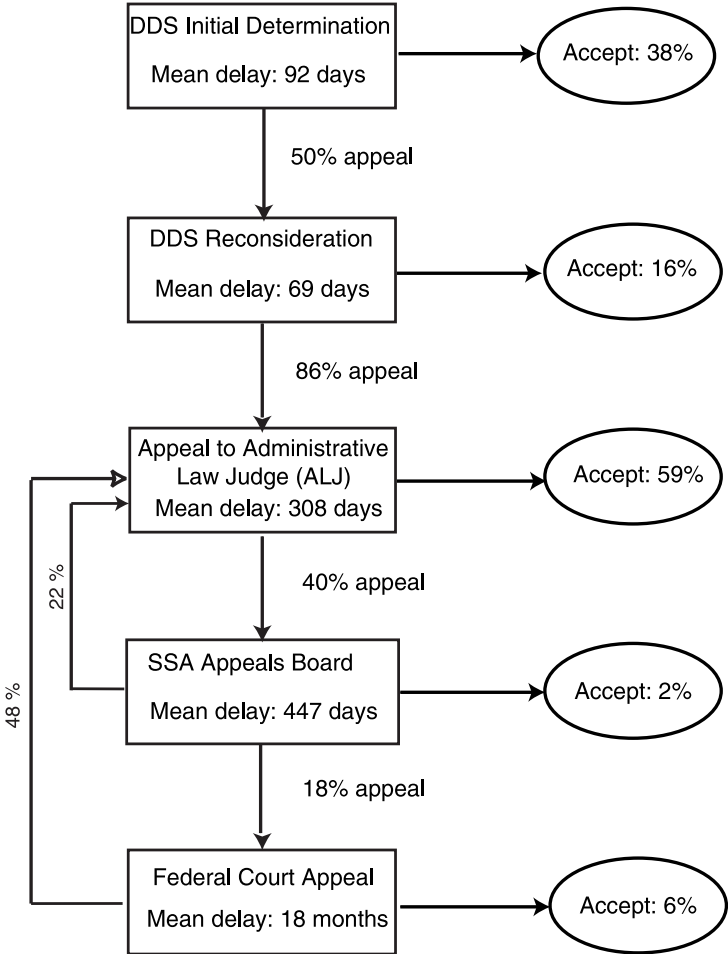
$$E [\tilde{a} - \tilde{d} | x] = 0 \quad (2)$$

- Note that unbiasedness implies that

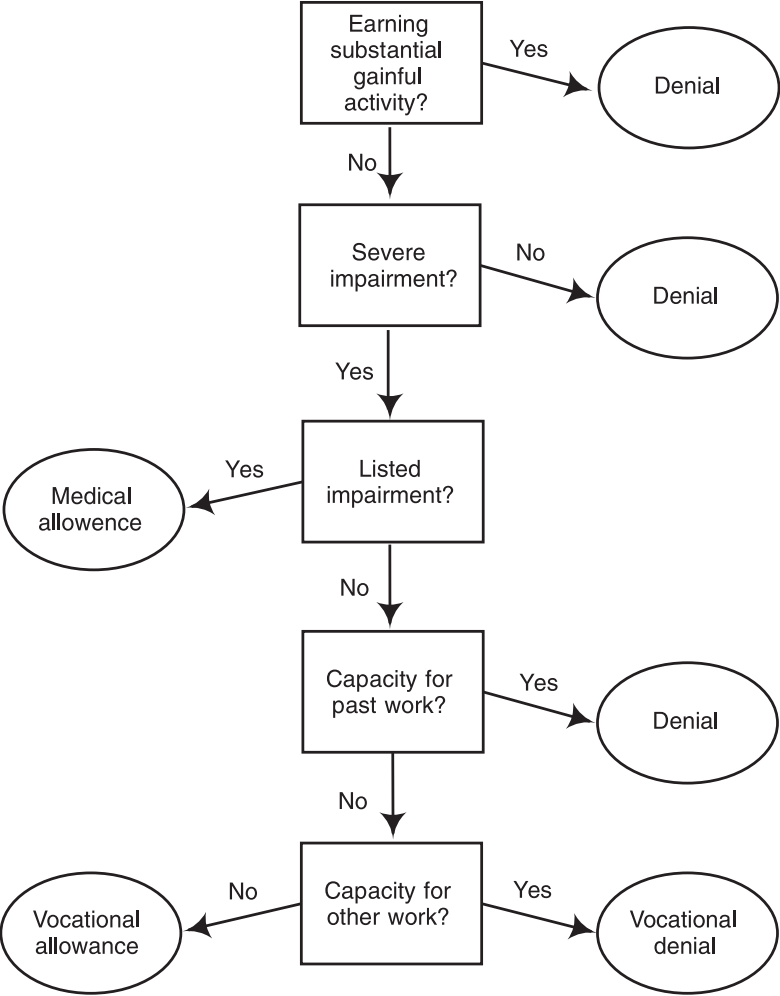
$$\Pr\{\tilde{a}|x\} = \Pr\{\tilde{d}|x\} \quad (3)$$

- We are unable to reject the unbiasedness hypothesis via GMM tests, OLS F-tests, the Bierens test, and the Horowitz-Spokoiny test
- Our failure to reject this hypothesis is not due to lack of power, since we decisively rejected the hypothesis when \tilde{a} is defined as the *first stage award decision* (i.e. by the DDS) instead of the *ultimate award decision*.

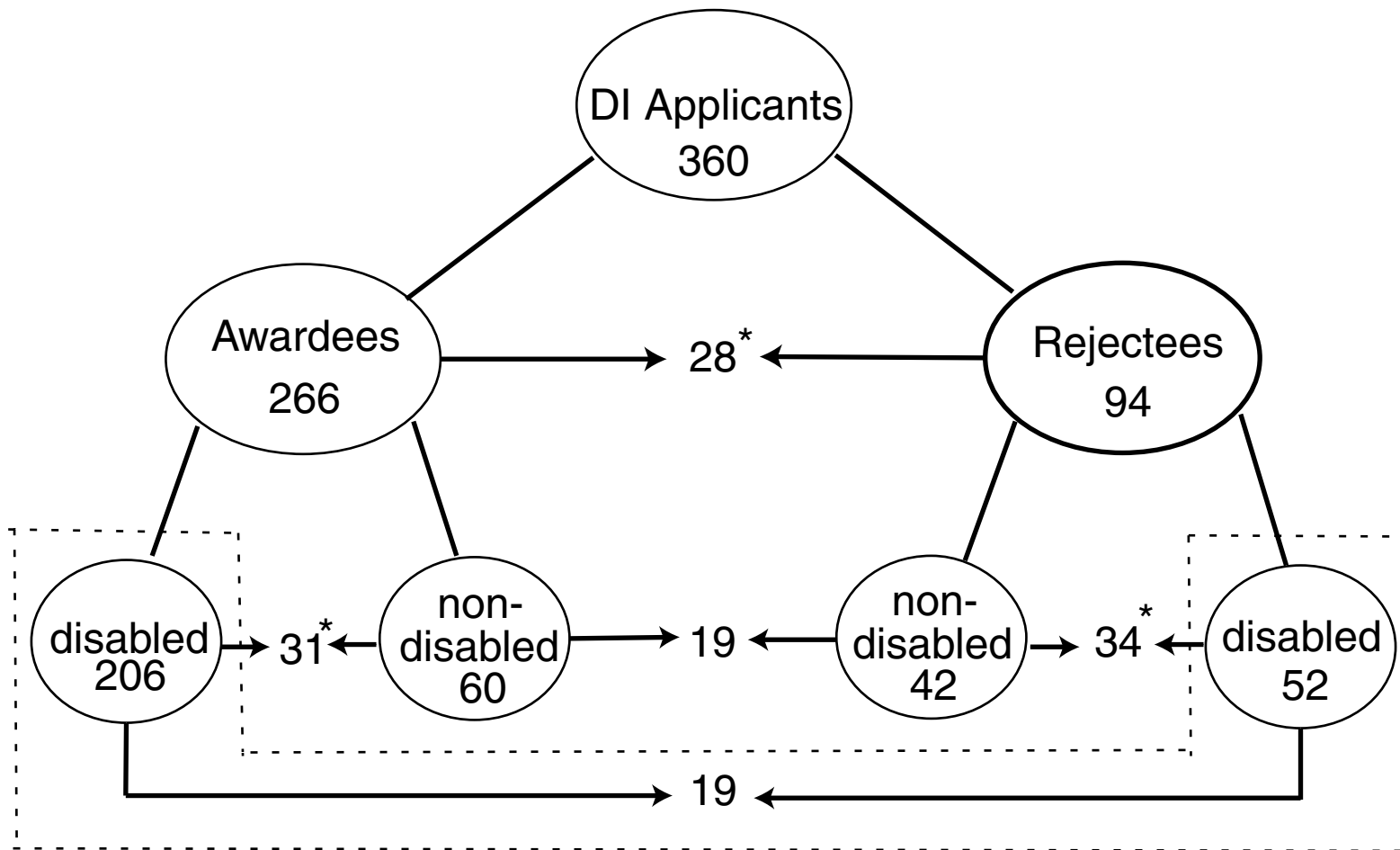
Summary of the SSI/SSDI Application and Appeal Process



The "5 Stage" decision process used by the DDS



Summary of Classification Errors in the DI Award Process



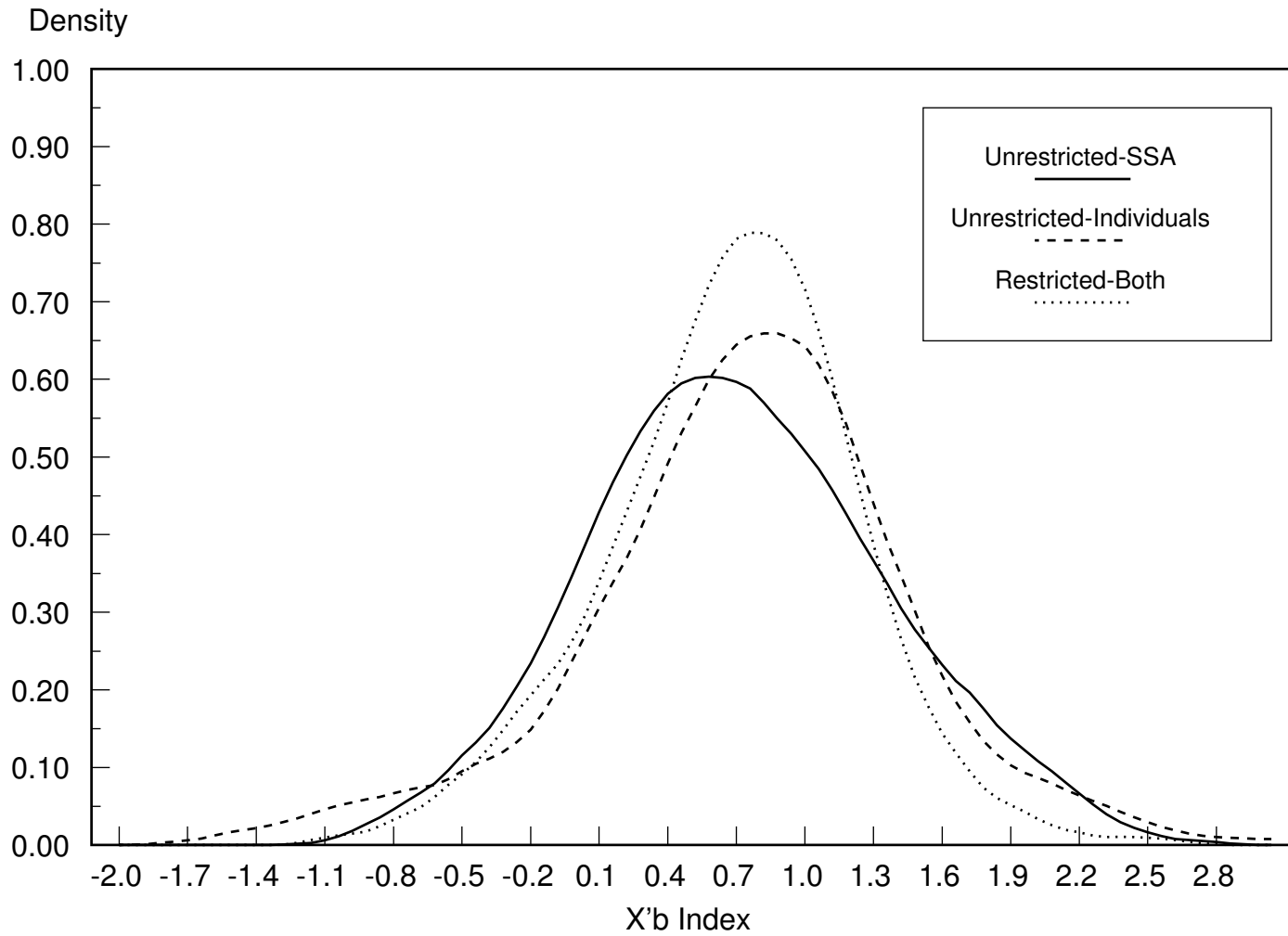
Testing Accuracy of Self-Reported Disability

- **Accuracy Test** Without loss of generality, we can represent both the individual's self-reported disability status and the SSA's ultimate award decision, and "true disability status" as "threshold rules"

$$\begin{aligned}\tilde{\tau} &= I\{x'\beta_{\tau} + \varepsilon_{\tau} \geq 0\} \\ \tilde{d} &= I\{x'\beta_d + \varepsilon_d \geq 0\} \\ \tilde{a} &= I\{x'\beta_a + \varepsilon_a \geq 0\}\end{aligned}\tag{4}$$

- **Accuracy Hypothesis** $\beta_d = \beta_a$
- Although $\tilde{\tau}$ is latent, both \tilde{d} and \tilde{a} are observable.
- We are unable to reject the hypothesis that $\beta_d = \beta_a$ when we assume that $(\varepsilon_d, \varepsilon_a)$ is bivariate normal with mean zero, unit variances, and unrestricted covariance ρ

Distribution of $X\beta_a$ and $X\beta_d$ scores



Rational Unbiased Reporting (RUR) Hypothesis

- **Assumption 1: (RUR Hypothesis)** $\beta_\tau = \beta_d = \beta_a$
- **Assumption 2:** $(\varepsilon_\tau, \varepsilon_d, \varepsilon_a)$ is trivariate normal with mean zero, unit variances (identification normalization) and

$$\varepsilon_a = \rho_a \varepsilon_\tau + v_a, \quad (5)$$

$$\varepsilon_d = \rho_d \varepsilon_\tau + v_d, \quad (6)$$

where v_a and v_d are independent of ε_τ and each other. Note that

$$\text{cov}(\varepsilon_d, \varepsilon_a) = \rho_a \rho_d \quad (7)$$

- **Assumption 3** $\rho_a = \rho_d \equiv \rho$. This implies that

$$\varepsilon = (\varepsilon_\tau, \varepsilon_a, \varepsilon_d) \sim \mathbf{N}(0, \Sigma), \quad (8)$$

where the covariance matrix Σ is given by

$$\Sigma = \begin{pmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho^2 \\ \rho & \rho^2 & 1 \end{pmatrix}. \quad (9)$$

Implications of (RUR) Hypothesis

- Assumptions 1-3 are sufficient to allow us to compute Bayes estimates of classification errors (and Type 1 and 2 error) of the Social Security disability award process under the weaker assumption that both \tilde{a} and \tilde{d} are noisy (but unbiased) indicators of “true disability status” $\tilde{\tau}$
- Recall that the award error rate can be written as

$$\Pr(\tilde{\tau} = 0 | \tilde{a} = 1) = \int \Pr(\tilde{\tau} = 0 | \tilde{a} = 1, x) f_X(x) dx, \quad (10)$$

where $f_X(x)$ is the density of observed characteristics of the applicant.

- This conditional probability can be rewritten as

$$\begin{aligned} \Pr(\tilde{\tau} = 0 | \tilde{a} = 1, x) &= \Pr(\tilde{d} = 0) \Pr(\tilde{\tau} = 0 | \tilde{a} = 1, \tilde{d} = 0, x) \\ &\quad + \Pr(\tilde{d} = 1) \Pr(\tilde{\tau} = 0 | \tilde{a} = 1, \tilde{d} = 1, x). \end{aligned} \quad (11)$$

Implications of (RUR) Hypothesis (continued)

- Now note that

$$\Pr\left(\tilde{\tau} = 0 \mid \tilde{a} = 1, \tilde{d} = 0, x\right) = \frac{\Pr\left(\tilde{\tau} = 0, \tilde{a} = 1, \tilde{d} = 0 \mid x\right)}{\Pr\left(\tilde{a} = 1, \tilde{d} = 0 \mid x\right)}, \quad (12)$$

and similarly for $\Pr\left(\tilde{\tau} = 0 \mid \tilde{a} = 1, \tilde{d} = 1, x\right)$.

- The probabilities in the numerator and the denominator in the expression above can be easily computed, given the trivariate normal distribution of ε using the GHK algorithm and the coefficient estimate for β from Benítez-Silva et al. (2001). In the example of the probability in the numerator of the above equation is given by

$$\Pr(\tilde{\tau} = 0, \tilde{a} = 1, \tilde{d} = 0 \mid x) = \Pr(\varepsilon_{\tau} < -x\beta, \varepsilon_a \geq -x\beta, \varepsilon_d < -x\beta \mid x). \quad (13)$$

- The computation of the *rejection probability* can be done in a similar manner.

Implications of (RUR) Hypothesis (continued)

- Calculation of Type 1 and 2 error rates. We have:

$$\Pr(\tilde{a} = 0 | \tilde{\tau} = 1) = \Pr(\tilde{\tau} = 1 | \tilde{a} = 0) \frac{\Pr(\tilde{a} = 0 | x)}{\Pr(\tilde{\tau} = 1 | x)}, \quad \text{and}$$
$$\Pr(\tilde{a} = 1 | \tilde{\tau} = 0) = \Pr(\tilde{\tau} = 0 | \tilde{a} = 1) \frac{\Pr(\tilde{a} = 1 | x)}{\Pr(\tilde{\tau} = 0 | x)}.$$

- **Award Error Rate:** $\Pr(\tilde{d} = 0 | \tilde{a} = 1) = .23$
- **Rejection Error Rate** $\Pr(\tilde{d} = 1 | \tilde{a} = 0) = .61$
- **Type I Error Rate** $\Pr(\tilde{a} = 0 | \tilde{d} = 1) = .23$
- **Type II Error Rate** $\Pr(\tilde{a} = 1 | \tilde{d} = 0) = .68$

Model	Error Type			
	Award	Rejection	Type I	Type II

a . Equicorrelation case: $\rho_{ad} = .12$; $\rho_{a\tau} = \rho_{d\tau} = \sqrt{\rho_{ad}}$

$\sigma_{\tau} = 1$ 21.71% 59.94% 23.71% 67.67%

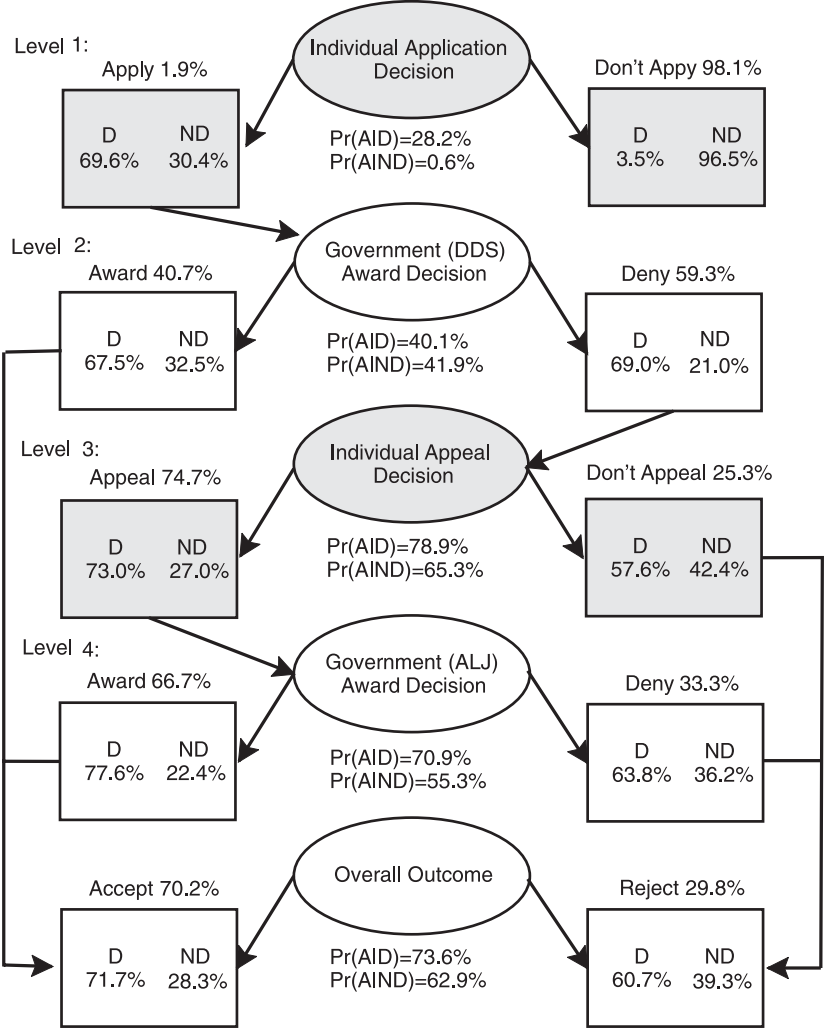
b . Asymmetric correlation case with $\rho_{ad} = .12$; $\rho_{a\tau} = .4$; $\rho_{d\tau} = .2$

$\sigma_{\tau} = .9$ 19.34% 57.49% 21.44% 55.05%

c . Asymmetric correlation case with $\rho_{ad} = .12$; $\rho_{a\tau} = .5$; $\rho_{d\tau} = .1$

$\sigma_{\tau} = .8$ 16.10% 52.03% 19.27% 41.60%

Analysis of Classification Errors in SSA's Disability Award Process



Can a Computerized Screening Rule Outperform the SSA?

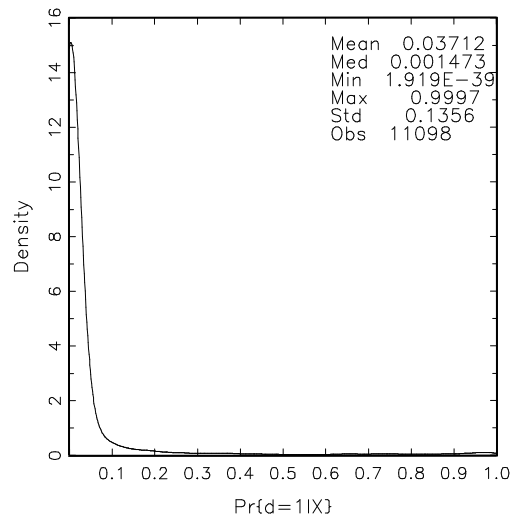
- Even if we believe that self-reported disability status \tilde{d} is truthfully reported in an anonymous survey such as HRS, SSDI/SSI applicants have a clear incentive to lie about their disability status to the SSA
- However we can use the HRS data to regress \tilde{d} against a vector x of “objective” health conditions and ADLs, such as “do you have heart problems?” “do you have diabetes?” “have you had a stroke?” etc.
- Using the HRS data we then can compute predicted probabilities $\hat{\Pr}(\tilde{d} = 1|x)$ that a person is truly disabled. These predicted probabilities depend only on the observable health/demographic characteristics x
- Define an **acceptance rule** of the form

$$\tilde{a} = I\{\hat{\Pr}(\tilde{d} = 1|x) \geq \lambda_c\} \quad (14)$$

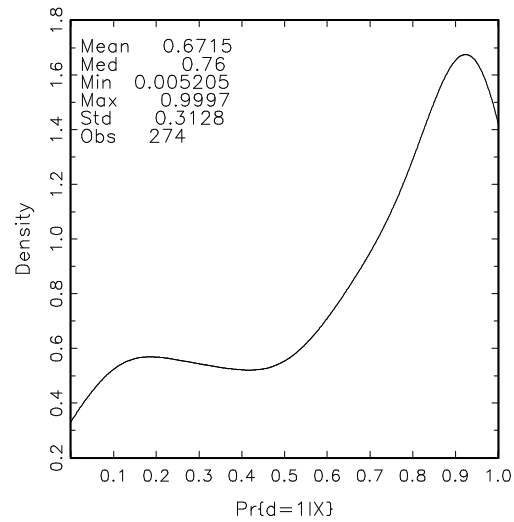
- By varying the cutoff λ_c we can achieve any desired award rate.

Distributions of $\Pr\{\tilde{d} = 1|x\}$ for Different Subsamples

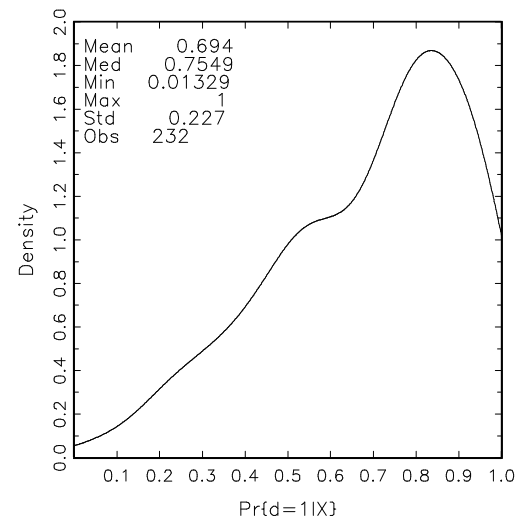
Distribution of $\Pr\{d=1|x\}$
Full Sample



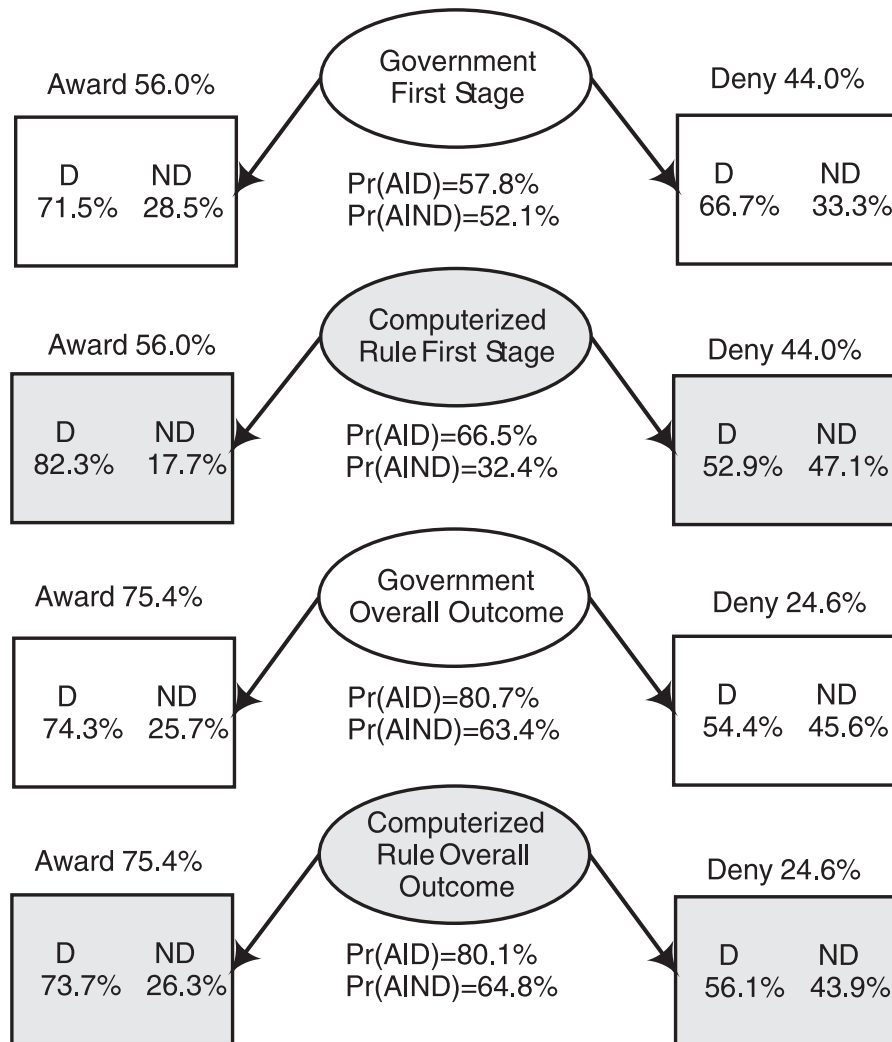
Distribution of $\Pr\{d=1|x\}$
Applicant Subsample



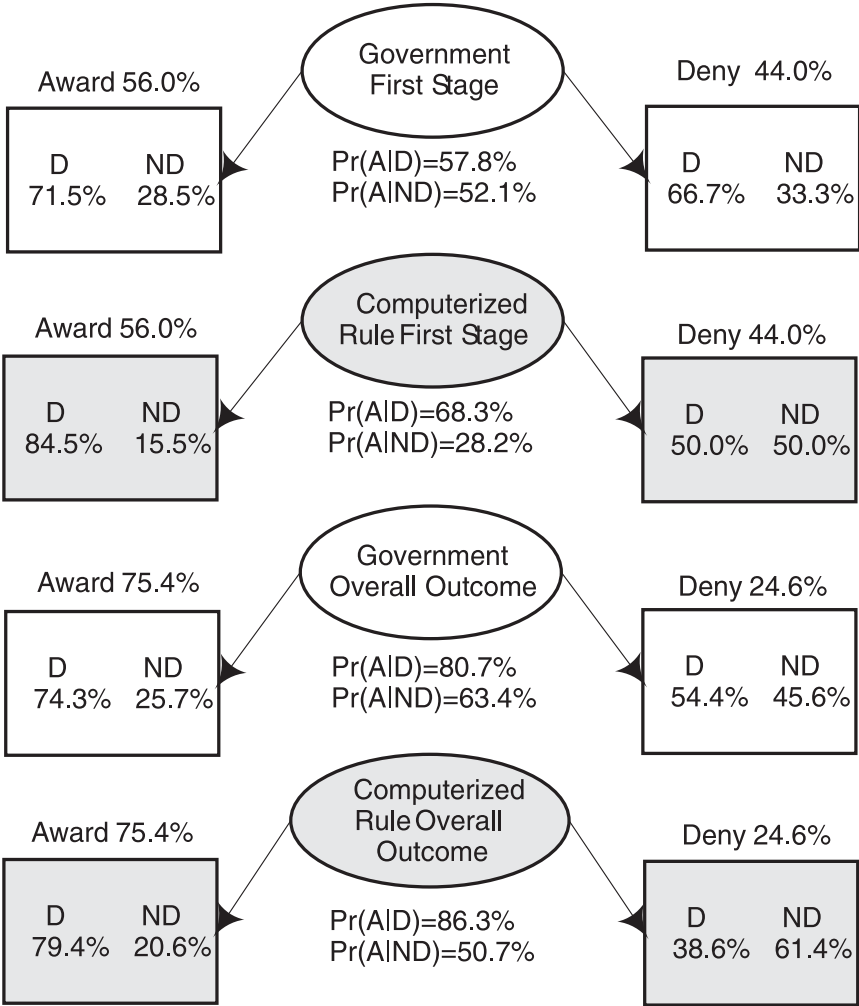
Distribution of $\Pr\{d=1|x\}$
Applicant Subsample



Computerized Screening Rule: Full Sample



Computerized Screening Rule: Subsample of SSI/SSDI Applicants



Conclusions

1. We have shown that the SSDI/SSI award process used by the SSA is very noisy, resulting in award error rates of over 20% and rejection error rates of over 50%
2. Most of the screening in the DI award process is done by the applicants themselves, via self-screening the application and appeal decisions, not by the SSA bureaucracy.
3. The first stage decisions by the DDSs have low award rates and extremely high rates of rejection error. They appear to be adopting a strategy of “when in doubt, reject”
4. Contrary to the GAO analysis, we find that the appeal stage to the ALJs, and the high rate of reversals, substantially reduces the rate of rejection errors without increasing the rate of award errors.

Conclusions (continued)

5. The ALJs are not necessarily better judges of disability than the DDSs: in particular their rate of rejection errors is nearly as high as the DDSs. ALJs benefit from the fact that they are judging a self-selected pool of rejectees who chose to appeal, and these individuals are more likely to be disabled.
6. Although some of our results depend on the assumption that self-reported disability is equal to “true disability” and the SSA ultimate award decision is a noisy indicator of true disability, we showed that our calculated classification and Type 1 and 2 error rates are hardly changed under the weaker assumption that both the individual self-report and the SSA’s ultimate award decisions are noisy indicators of true disability.

Conclusions (continued)

7. We compared the performance of the SSA's award process to an alternative computerized screening rule that awards benefits to an applicant if the applicant's predicted probability of being disabled (as a function of "objective" health and demographic characteristics) is sufficiently high. We find that the computerized screening rule results in significantly lower classification error rates,
 - 7-A. At the DDS stage, award error rates are nearly cut in half, from 29% to 16%, and rejection error rates fall from 67% to 50%
 - 7-B. If the computerized rule were to replace the overall award process, award error rates would fall from 26% to 21% and rejection error rates would fall from 54% to 39%

Caveats

- In previous work we have described the long delays at various stages of the application and appeal process.
- These delays act as a significant barrier to applications and appeals.
- Essentially the delays can be viewed as **type-dependent application fees** since the opportunity cost of delays in application and appeal is presumably near zero for someone who is truly disabled, but can be significant for someone who is not disabled.
- If a computerized rule lead to a significant speedup in the award and appeal process, it could induce additional applications and reduce the “signalling cost” among individuals who are not truly disabled.
- Theoretical work by Don Parson’s shows that application fees can significantly improve the operation of the SSDI/SSI award process, improving its ability to sort even when there is substantial inherent noise in the decision making process.

What We Recommend

- Replace the DDS stage by a computerized screening rule, but retain human decision makers in the appeal stage
- Offset the faster time to decision of a computerized rule by an explicit financial application fee
- This application fee would cover the costs of a medical exam, *by doctors who are agents of the SSA, not of the DI applicant*
- Further work needs to be done, using a structural model of the application and award process, to fine tune just how high the application fee needs to be in order not to result in induced entry as a result of speeding up the decision process.
- Estimates of $\hat{\Pr}(\tilde{d}|x)$ could be obtained by doing a Nagi-style study where independent teams of human experts make a moderated group decision about whether particular applicants are disabled or not.

What the National Academy of Sciences Recommended

- *The Dynamics of Disability* (2002) report by the Institute of Medicine and the National Research Council
- “In 1996 SSA requested that the Institute of Medicine (IOM), in collaboration with the National Research Council’s Committee on National Statistics (CNSTAT) conduct an independent review of the statistical design and content of the disability survey under development (National Survey of Health and Activity (NSHA)) and of its research plan for the redesign of the disability decision process.” (p. 2)

What the National Academy of Sciences Recommended

- In the chapter on “Improving the Disability Decision Process” the study noted that “The committee then reviewed the general features and directions specified by SSA in its research plan and the individual projects within the plan with reference to each of the research steps identified in the framework. . . . It identified critical elements of a research design that were missing from the SSA’s plan, expressed serious concerns about these gaps, and made recommendations for redirected and new research effort.” (p. 120)
- *“It urged SSA to adopt a rigorous research design process and to develop, early in the research, objective validation criteria and validation plans to be able to make the ultimate judgements on whether or not the proposed changes would yield the desired results. In issuing the interim report the committee hoped that the recommendations embodied in that report would be incorporated in the contract research that was under way and in new research not yet initiated at the time.”* (p. 123)

How SSA Responded to these Recommendations

“However, after the committee issued its second interim report (IOM, 1998), SSA undertook an internal reevaluation of its disability decision process redesign initiatives, SSA concurred with several of the committee’s conclusions and some of the recommendations. However, rather than undertaking the additional research and redirection of the research as recommended by the committee, SSA decided to no longer actively pursue the new decision-making process proposed in the Disability Redesign, but to improve the current process, focusing at this time on updating the Listings.” (p. 123)

Bottom Line of the NAS Report

“In conclusion, although SSA has deferred major redesign of the disability decision process, the committee believes that it is paramount that the determination of disability not only be timely, understandable, straightforward, and feasible, but also provide accurate and consistent decisions that are fair to the claimant and the government. To this end the committee believes that SSA should undertake a systematic, long term program of research — intramural and extramural — that provides baseline information on all the key aspects of the current disability decision process and subsequent evaluative data on all future change aimed at improving the effectiveness and efficiency of the work disability determination process currently in use in the United States.” (p. 139)

Comments on My Experience with SSA

- **SSA is massively underspending on essential R&D**
- **SSA is completely unequipped** in terms of knowledge and ability to develop or use models that could enable it to advise Congress and the President on reforms currently being considered by Congress
- **Main Cause of the Problem** Narrow minded attitudes of political appointees “running the shop” and the “start/stop” and short-run mentality due to turnover of successive Commissioners
- **My recommendation:** SSA should be a quasi-independent government agency, similar to the Fed, with Commissioners who have long terms not linked to the 4 year election cycle
- **Is there any hope?** Competition from other agencies that have better leadership and a more research oriented culture (e.g. the CBO) may create pressure for SSA to “get its act together”

How Much R&D Should SSA be Doing?

- US is not even on the top 5 internationally in terms of R&D spending as a fraction of GDP (public and private R&D)
- US government has high R&D spending, and a rapid rate of increase (growing 4.8% from 2004 to \$132 billion in FY 2005), but 80% of that increase will go to defense (e.g. “Star Wars” and nuclear weapon testing).
- US spends an order of magnitude more on R&D and programming of video games for children (e.g “The Sims”), than it does on software for modeling and simulating fiscal policy and Social Security.
- Overall, 57% of US government R&D spending is defense-related.
- To his credit, President Bush nearly doubled the budgets of NIH and NSF during his first term. But now in time of huge budget deficits, their budgets have stopped growing.

How Much R&D Should SSA be Doing?

- Taking the US rate of spending (public and private) as a fraction of GDP, 2.6%, times the 5% of GDP that is devoted to Social Security, SSA would have to spend *\$14 billion per year on R&D if it were to spend at the national average.*
- In reality, *SSA spends less than \$100 million per year on R&D* In comparison, the annual budget for NIH is \$28 billion and NSF's budget is \$5.4 billion.
- I don't want to sound like another special interest, clamoring for my "cut" of the Federal dollar, and we need to recognize that the US government is in very poor fiscal shape and cuts need to be made somewhere
- However it is appalling that SSA does not even seem to be aware, or capable of understanding/utilizing the existing "free" research results financed by NSF and NIH.
- **Conclusion:** SSA is not "pulling its weight" and the US will ultimately pay a big price for the narrow mindedness of its leadership.

Using a Life-Cycle Model to Predict the Impact of the “\$2 for \$1 offset proposal” in the 1999 Ticket to Work and Work Incentives Improvement Act (TWWIIA)

John Rust, *University of Maryland*

Hugo Benítez-Silva, *SUNY-Stony Brook*

Moshe Buchinsky, *UCLA, NBER, and CREST-INSEE*

Paper Originated When I was an Advisor to SSA

- Written as an “executive summary” for higher ups in the SSA when I was working as an advisor on the design of large scale controlled experiment (“demonstration project”) intended to assess the \$2 for \$1 offset, that SSA was mandated to conduct under the TWWIIA act.
- The \$2 for \$1 offset is essentially a reduction in the tax on labor supply by SSDI beneficiaries from 100% to 50%, intended as an incentive to get healthier DI beneficiaries to return to work and leave the roles.
- Problem: this reduction in the effective tax rate makes the proposed policy change strictly preferable to the *status quo*. Congress, in its wisdom, was concerned about the possibility of *induced entry* by applicants on the margin between applying and not applying.
- Forecasts by the SSA and CBO predicted that the induced entry effect could be significant and would result in a significant increase in the net cost of the SSDI program.

The demonstration project

- Congress mandated a demonstration project to assess the magnitude of the induced entry effect *before* actually implementing the \$2 for \$1 offset.
- Unfortunately, a panel of advisors lead by Nancy Tuma of Stanford, concluded it would be prohibitively costly and time consuming to assess the induced entry effect via a randomized experiment, and there is no guarantee that the results would be reliable.
- Rather than simply throw in the towel, I wrote this report (*pro bono*) to show how the life cycle model could be adapted, at very modest cost, to provide forecasts of the induced entry effect and the budgetary and welfare impacts of the \$2 for \$1 offset.
- Unfortunately, this report has been largely ignored and the TWWIIA panel has been disbanded, and the demonstration project and associated planned survey data collection (National Study on Health and Activity, NSHA) were cancelled by the SSA Commissioner JoAnne Barnhardt (a Bush appointee).

Why was the demonstration project cancelled?

- Seems like an excellent opportunity to collect data. Rarely have the ability to conduct controlled experiments on real human subjects
- Barnhardt's reasoning, reportedly, was that she wanted results "better, cheaper, faster"
- One way to get things "better, cheaper, faster" is not to do it all.
- However hard to see how SSA can be complying with the 1999 law which *mandates* that a controlled experiment be done.
- Budget constraints are a reality, and experiments are very time-consuming and costly. I can relate to that. But it is hard to see then why SSA would also ignore any attempt to comply with the law by using alternative methodologies, including models, especially when these are available to it at much lower cost (If they would have accepted my report, it would have been at zero cost).

What Issues/Initiatives is SSA Devoting its Scarce Resource to?

- **AeDiB** Accelerated Electronic Disability Claims processing initiative. Accelerates a previous (stalled) effort called eDiB.
- Awarded \$900 million contract to IBM to develop AeDiB.
- \$600 million of the amount is devoted to *training* of SSA employees. With 15,000 employees, *this amounts to a training cost of \$40,000 per employee!*
- In view of the (outrageously) high cost of this software, SSA should change the acronym from **AeDiB** to **PayIBM**
- Hopefully, after spending this much SSA will not end up with software that does not work and has to be scrapped, as has happened with large software projects at IRS, FAA, and FBI.

The politics of policymaking

- Unfortunately, there is a very limited role for economic science in policymaking under the Bush administration.
- Policy making is all about politics: economic advisors are basically “hired guns” and “cheerleaders”. Tell the administration something it does not want to hear and you’re fired.
- “Intelligence failures” are ultimately due to lack intelligence in the White House: the management style of George Bush that does not tolerate, much less encourage, informed debate and the best scientific advice on the issues.

**Paul O’Neill, Treasury secretary fired by Bush due to unwillingness to pay
“The Price of Loyalty”**

“I realized that it’s very hard for an organization or institution to achieve more than the leader can imagine. If you determine to run a five minute mile, you’ll never run a four minute mile. The leader sets the conditions as to what it is that we aspire to. It’s not clear to me that you can create a process that will impact whatever the leader brings with him, in terms of his instinct of imagining what could be. I’d like to think that maybe that’s not right and think that someone can be a leader and see or smell a new idea and then own it. There has to be an openness of wanting to do that. There has to be a market there, otherwise you can just spin your wheels all you want to. You can have a lot of great ideas, and if the store is closed, it’s not going to make any difference . . .”

O’Neill quoted p. 293 from R. Susskind, *The Price of Loyalty*

The ugly story of how the Medicare Bill was passed

- A complex 2,000+ page bill drafted in the “backrooms” by Republicans with the assistance of the pharmaceutical companies. Democratic congressmen and senators were prevented from seeing the bill until hours before it was voted, at 4am in the morning.
- White House failed to disclose material information, including its own internal estimate of the cost of bill which, at \$534 billion over 10 years, was 33% higher than the figures Congress had been given.
- Tommy Thompson and Dick Cheney on the House floor twisting arms of “swing vote” conservative Congressmen, offering them “bribes” and threats to get them to capitulate.

The ugly story of how the Medicare Bill was passed

- Thomas Scully, head of CMS, threatened to fire underlings who disclosed their (high) cost estimates of the bill. Scully steps down after passage of the bill and takes a high \$ job as lobbyist to pharmaceutical companies, and was subsequently “pardoned” from conviction for potential ethics violations by Tommy Thompson.
- Billy Tauzin, Democratic congressman from Louisiana, retires from Congress to receive a \$2.5M/year pharmaceutical lobbying job as a payoff for his efforts to get the bill passed.
- No surprise that the bill prevents the U.S. government to use any leverage to bargain with pharmaceutical companies to keep prices down.

Cheer up! Things could be worse!

- Consider Brazil. Social Security in Brazil is so underdeveloped that they do not even have the equivalent of Social Security numbers!
- Without any reliable way to verify eligibility, social security fraud is rampant in Brazil. There are towns in the Amazon where 70% of the population is receiving Social Security!
- Revenue collection is just as backward. No reliable way to enforce individuals and firms that don't pay social security taxes, nor even a reliable record-keeping system to determine if they have paid or not.
- Reformers who try to expose fraud in the system are subject to death threats.
- Example, most recent Minister, Amir Lando, sacked by Brazil's president in March after visiting US (University of Maryland) to seek help from academics to help bring IT to bear to solve Brazil's huge social security problems.

Also, things have been worse in the US in the past

Richard Bellman's account of why he invented the term “dynamic programming”

“An interesting question is, ‘Where did the name, dynamic programming, come from?’ The 1950s were not good years for mathematical research. We had a very interesting gentleman in Washington named Wilson. He was Secretary of Defense, and he actually had a pathological fear and hatred of the word, research. I’m not using the term lightly; I’m using it precisely. His face would suffuse, he would turn red, and he would get violent if people use the term, research, in his presence. You can imagine how he felt, then, about the term, mathematical. The RAND Corporation was employed by the Air Force, and the Air Force had Wilson as its boss, essentially.”

Hence, I felt I had to do something to shield Wilson and the Air Force from the fact that I was really doing mathematics inside the RAND Corporation. What title, what name, could I choose? In the first place I was interested in planning, in decision making, in thinking. But planning, is not a good word for various reasons. I decide therefore to use the word, 'programming.' I wanted to get across the idea that this was dynamic, this was multistage, this was time-varying — I thought, let's kill two birds with one stone. Let's take a word that has an absolutely precise meaning, namely dynamic, in the classical physics sense. It also has a very interesting property as an adjective, and that is it's impossible to use the word, dynamic, in a perjorative sense. Try thinking of some combination that will possibly give it a perjorative meaning. It's impossible. thus, I thought dynamic programming was a good name. It was something not even a Congressman could object to.”

(quoted from p. 159 of **Richard Bellman's** biography, *The Eye of the Storm* World Scientific Publishing Company, 1984).

Finally, despite my current cynicism, I remain optimistic

- I believe that dynamic programming can be a useful practical tool, but not for the government, at least in the short run.
- DP/mechanism design methods *might* be useful for firms, as a management tool to help them take better decisions and design more efficient contracts
- My paper, “Impact of Retiree Health Care on Faculty Retirement Decisions” is a humble start down this path.
- The paper develops a special version of the life cycle model relevant to the case of a *tenured faculty member, who is approaching retirement.*
- I use the model to compare several different strategies for dealing with the high cost of health care, particularly *retiree health care.*

The problem of out of control health care costs

- Unlike Social Security Reform, *a second order problem facing the US*, health care reform is *a first order problem that America must confront very soon!*
- The problem is ubiquitous, and nobody can escape it, *not even the US military!*
- An articles in the *New York Times* business section, Thursday April 14th, “A New Call to Arms: Military Health Care” describes the main military health care plan, *Tricare* whose cost has doubled since 2001 and is expected to cost *at least \$100 billion over the next decade!*
- GM is being crushed by health care costs: currently health care costs account for \$1500 of the total cost of producing each vehicle, compared to \$200 for Japanese cars. GM’s retiree health care costs alone cost 5.3 billion last year, up 20% from the previous year. For each of its workers, GM has 2.5 retirees that it is obligated to pay (generous) retiree benefits to!

The problem of out of control health care costs

“Tricare for Life is one of a long list of assurances, like prescription drug benefits for the elderly, that Washington is making to American citizens at a rate of more than \$1 trillion a month. The government’s unpaid for promises grew by more than \$13 trillion last year, a sum larger than the nation’s 2004 economic output, and they now surpass \$43 trillion, said David A. Walker, comptroller general of the United States.”

quoted from page C7 of *New York Times* article, “Another Call to Arms: Health Care for the Military”, Thursday, April 14, 2005.

Results: Costs of Alternative Compensation Plans

Panel A: Expected Discounted Values (at 2%) (amounts in thousands)					
Compensation Item	Base Case	Case 0	Case 1	Case 2	Case 3
Wages	1047.17	1497.40	1370.9	1222.48	1183.06
Pension Contributions	52.36	74.87	68.6	61.12	59.15
Employee Health Insurance	537.55	863.41	764.3	303.43	294.82
Retiree Health Insurance	553.13	113.64	0.0	33.79	42.40
Total Compensation	2190.21	2549.31	2203.81	1620.83	1579.43

Panel B: Breakdown as a Percent of Total Expected Discounted Compensation					
Compensation Item	Base Case	Case 0	Case 1	Case 2	Case 3
Wages	47.8	58.7	62.2	75.4	74.9
Pension Contributions	2.4	2.9	3.1	3.8	3.8
Employee Health Insurance	24.5	33.9	34.7	18.7	18.7
Retiree Health Insurance	25.3	4.5	0.0	2.1	2.7
Total Compensation	100.0	100.0	100.0	100.0	100.0