RESEARCH PAPER NO. 9701

HOUSING BIAS IN THE CPI AND ITS EFFECTS ON THE BUDGET DEFICIT AND THE SOCIAL SECURITY TRUST FUND

by

James E. Duggan
Office of Economic Policy
U. S. Department of the Treasury,

Robert Gillingham
Fiscal Affairs Department
International Monetary Fund

and

John S. Greenlees
Office of Prices and Living Conditions
U. S. Bureau of Labor Statistics

December 1999

The work for this paper was initiated while all three authors were with the Department of Treasury. Views expressed are those of the authors and do not necessarily represent the policies or views of the Department of the Treasury, the International Monetary Fund or the Bureau of Labor Statistics. The authors thank Bing Chen and David Wilcox for comments on an earlier version of this paper.

HOUSING BIAS IN THE CPI AND ITS EFFECTS ON THE BUDGET DEFICIT AND THE SOCIAL SECURITY TRUST FUND

Potential bias in the Consumer Price Index (CPI) and its effect on the budget deficit have been the subject of intense debate over the past several years. Alan Greenspan, Chairman of the Federal Reserve Board of Governors, testified before the Senate Finance Committee to his belief that the CPI has an upward bias of between 0.5 and 1.5 percentage points per year and noted that a reduction of one percentage point in the indexation of tax rates and benefit programs would save the federal government roughly $150 billion over five years (Greenspan, 1995). The Finance Committee subsequently appointed the Advisory Commission to Study the Consumer Price Index (known as the Boskin Commission), which estimated that the CPI overstates prospective changes in the cost of living by more than 1.1% per year, and called this a substantial improvement over much larger historical biases (Advisory Committee, 1996). There is not, however, consensus on even the direction of bias, much less the size. Empirical evidence is sketchy, and conclusions about bias abstract from the fact that the major benefit program indexed by the CPI—Social Security—uses an index that does not even include Social Security beneficiaries in its coverage. In this light, it is hard to make *definitive* estimates of bias.

We do, however, have an empirical estimate of a major, *known* bias in the CPI. Prior to 1983 for the Consumer Price Index for all Urban Consumers (CPI-U) and to 1985 for the Consumer Price Index for Urban Wage Earners and Clerical Workers (CPI-W), the homeownership component was based—very roughly—on the costs incurred in purchasing, financing and maintaining a house. More specifically, the home-purchase expenditure weight was the net purchase of owner-occupied houses in the survey period, and the mortgage-interest expenditure weight was the total interest (undiscounted) that would be paid over half the term on all mortgages incurred during the survey period. These expenditure weights were indexed over time by a home-purchase price index and the product of a home-purchase price index and a mortgage-interest rate index, respectively. This methodology was without conceptual foundation (see Gillingham (1980, 1983)), and its use resulted in a substantial upward bias in the CPI.

The purpose of this paper is to provide a more precise quantification of the effect of the housing bias in the CPI on both the federal budget and the Old Age and Survivors Insurance and Disability Insurance (OASDI) Trust Funds using the Continuous Work History Sample (CWHS) and the Master Beneficiary Record (MBR) files from the Social Security Administration (SSA). The CWHS is an earnings history sample for 1 percent of participants in the Social Security system. It contains almost 2.7 million records with actual earnings histories spanning the period 1951 to 1992. The MBR is a benefit history file that contains associated OASDI benefits for the CWHS records. The combined CWHS–MBR file provides a comprehensive basis for measuring the effect of the housing bias in the CPI.

The remainder of the paper is organized as follows: Section 1 recounts the history of indexation of Social Security benefits; section 2 specifies precisely the role of the housing bias in past indexation; section 3 estimates the effects of this bias on the Federal budget and the OASDI Trust Funds; section 4 estimates the effects on individual birth cohorts; and section 5 concludes.

1. INDEXATION IN SOCIAL SECURITY

In order to put the CPI-bias issue in perspective, we describe briefly the role of indexation in the Social Security program. Four time periods are relevant to this topic: 1979 to present, 1973 to 1978, 1968 to 1972, and pre-1968. The distinction between initial retirement and post-retirement benefits is also important. Across all time periods, the computation of Social Security benefits at initial entitlement comprises three fundamental steps: (i) determining the average monthly earnings in covered employment, (ii) calculation of the benefit (referred to as the primary insurance amount or PIA) by applying a progressive rate formula to the average earnings, and (iii) adjusting the PIA for early or delayed retirement and spousal benefits.

Under current-law procedures, in effect since 1979 pursuant to the 1977 Social Security Amendments, prior covered earnings, the maximum contribution base, and the benefit formula brackets are adjusted automatically by an economy-wide wage index. Post-entitlement benefits, beginning at age 62 for retirees and any age for disability beneficiaries, are adjusted each year by the CPI-W.

Post-entitlement benefits under the rules applicable between 1973 and 1978 were also adjusted by changes in the CPI-W. The major difference from current law was in the calculation of initial benefits. From 1973 to 1978, earnings histories were not indexed in calculating the average monthly earnings, but the marginal benefit *rates* use to calculate the PIA were adjusted by changes in the CPI-W. These features, along with wage indexation of the contribution base, became "automatic" with the 1972 Social Security Amendments (modified by two Amendments in December 1973). The indexation of marginal benefit rates also created potential financing problems for the program. As long as the index used to adjust benefit rates is increasing, the average replacement rate—the ratio of initial benefits to final wages—will eventually increase. Given the levels of wage and price growth that prevailed in the mid-70s, replacement rates actually increased very rapidly under this scheme. This "overindexation" — which existed independent of any bias in the CPI—was corrected by the 1977 Amendments (effective for retirees born after 1916), which changed the way initial benefits are calculated so that replacement rates would stabilize. As we demonstrate below, the CPI bias has its largest impact on pre-1977 law beneficiaries, since it caused an upward drift in initial replacement rates.

Though automatic indexation of benefits was not implemented until June 1975, movements in the CPI-W became the guides for benefit changes subsequent to the 1972-1973 Amendments. The 11% increase granted in June 1974 reflected an attempt to approximate the CPI change between September 1972, the date of the last increase, and June 1974 (Myers, 1993). Between 1968 and 1972, however, Congress legislated four *ad hoc* benefit increases, in 1968 (13%), 1970 (15%), 1971 (10%) and 1972 (20%), that expanded real benefits dramatically. The last three of these increases, in particular, bore no relationship to changes in the CPI-W. This contrasts with earlier and later changes that generally reflected changes in the CPI-W.

Table 1 illustrates the cumulative real benefit increases granted since 1955, the year following the Social Security Act that established the foundation for all subsequent changes in the benefit formula. Through 1968, benefit changes kept rough pace with the CPI and the replacement rate remained fairly stable. However, the next three adjustments resulted in substantial real benefit increases and the average replacement rate rose 27 percent in three years. These developments had an important influence on the major change to benefit calculation introduced in the 1977 Amendments.

2. ERRORS IN THE MEASUREMENT OF HOUSING COSTS IN THE CPI-W

The discussion above pointed out the important role of the CPI-W in Social Security benefit determination. Any error in measurement in the CPI-W could be translated into a significant error in benefit payments. An upward bias, for example, would imply that benefits would be raised more than intended, that is, more than what is necessary in order to maintain the purchasing power of benefits. As discussed earlier, prior to 1985 the CPI-W suffered the defect of an inappropriate treatment of homeownership costs. The consequence of this defect was a (generally upward) bias in the price index used to adjust benefits, particularly during the late 1970s and early 1980s when interest rates were unusually high. As a result, benefits were higher than they would have been in the absence of the homeownership bias and, as shown below, for some retirement cohorts the "overpayment" was substantial.

TABLE 1. SOCIAL SECURITY BENEFIT AND REPLACEMENT RATES, 1955-1972*

	Cumulative Changes Since 1955			
Year	**Benefits**	**CPI-W**	**Real Benefits**	**Replacement Rate**
1959	7.0	7.6	-0.5	34.2
1965	14.5	16.2	-1.5	33.5
1968	29.4	26.9	2.0	36.3
1970	48.8	41.2	5.4	40.3
1971	63.7	48.3	10.4	43.0
1972	96.4	53.6	27.9	51.2

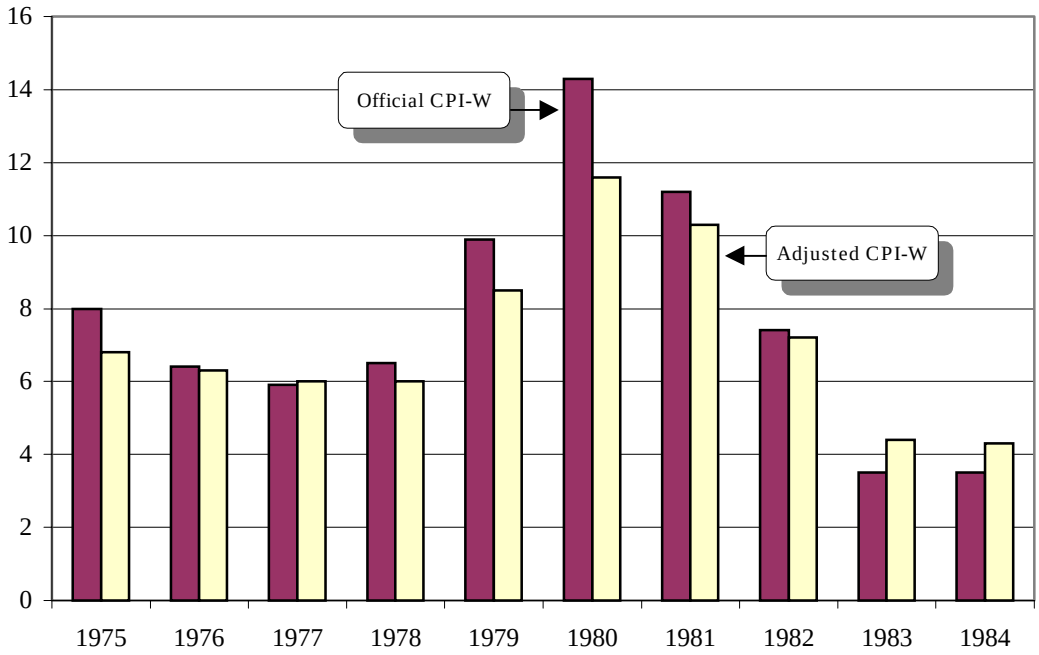
*Sources: Social Security Bulletin, Annual Statistical Supplement (1991); Replacement rates from Myers (1993, p. 363).

The conceptually appropriate rental-equivalence measure of homeownership was not incorporated into the CPI-W until 1985. Fortunately, the Bureau of Labor Statistics calculated an experimental version of the Consumer Price Index for all Urban Consumers (CPI-U) that substituted the residential rent index for the traditional—and flawed—homeownership cost index. This experimental index, denoted the CPI-U, X-1, was calculated back to 1967. We adjusted the CPI-W for mismeasurement of housing costs by assuming that it suffered from an identical housing bias as the CPI-U, and used this adjusted CPI-W to estimate the overpayment of Social Security benefits due to the homeownership bias in the CPI-W.[1]

[1]The CPI-U, X-1 differs from the CPI-W not only in its treatment of homeownership costs but also in population coverage, since the CPI-W applies only to wage-earner and clerical-worker households. However, the error caused by assuming similar housing biases is almost certainly relatively minor. Since the rental equivalence measure was incorporated in the CPI-U in 1983, two years prior to its incorporation in the CPI-W, we used the difference between housing cost inflation in the CPI-U and CPI-W as the measure of bias for these two years. Finally, we recognize that the adjustments made to the CPI-U to construct the CPI-U, X-1 are also subject to error. The rental equivalence measure incorporated in the CPI-W beginning in 1985 also used the CPI residential rent data, but in a way that more accurately reflected the different characteristics of owner-occupied and rental-occupied housing units. Moreover, BLS has made several subsequent changes to the CPI rental equivalence estimator to

Figure 1 charts the annual changes in the official CPI-W and our adjusted series over the period 1975 to 1984.[2] Three subperiods are noteworthy. The first occurred when the CPI-W increased by 8.0 percent between the second quarter of 1974 and the first quarter of 1975, the period used by SSA to calculate the benefit increase for 1975. Over the same period, the adjusted CPI-W increased by 6.8 percent. The second, and by far the most important subperiod, occurred from 1979 to 1982, when the CPI-W increased by a cumulative 60.0 percent while the adjusted series rose by 51.8 percent. Finally, in 1983 and 1984, the reverse situation occurred. Between those two years, the adjusted series increased by 8.9 percent, and the CPI-W rose by 7.2 percent.

FIGURE 1. PERCENT CHANGES IN THE CPI-W, WITH AND WITHOUT HOUSING CORRECTION, 1975–84



The manner in which a bias in the CPI affects beneficiaries differs fundamentally between the pre-1977 and post-1977 law (pre- and post-1979 benefit rules), though a significant effect can occur in either case. Table 2 shows cumulative percentage differences between the official CPI-W and the CPI-W adjusted for housing bias over the period 1974 to 1984. The first column is relevant to retirees under the pre-1977 (birth year < 1917) regime, disabled beneficiaries who first became eligible for benefits prior to 1979, and auxiliary beneficiaries who became eligible

---

improve accuracy. The deficiencies of the CPI-U, X-1 are minor, however, compared to those of the pre-1985 CPI homeownership measure.

[2]The percent changes in Figure 1 correspond to the periods used for determination of actual Social Security benefit increases. For 1975, the benefit increase was based on the change in the CPI-W between the second quarter of 1974 and the first quarter of 1975. Between 1976 and 1983, benefit increases were based on first quarter to first quarter increases in the CPI-W and, beginning with 1984, they were based on third quarter to third quarter changes in the CPI-W.

on a pre-1977 primary beneficiary account. Because the benefit formula was price-indexed each year after 1975, the full cumulative bias applies to all beneficiaries under the pre-1977 regime, *regardless of the year benefit receipt commenced*. Thus, individuals who retired in 1975 (or any other year under pre-1977 rules) received Social Security benefits that, by 1982, were 6.5 percent higher than what would have been experienced if the adjusted CPI-W were the indexing series and 4.7 percent higher by 1984 (assuming benefit receipt continued through 1984).[3]

TABLE 2. CUMULATIVE PERCENTAGE HOUSING BIAS, 1975–1984

	Cumulative Percentage Bias from:						
Cumulative To:	1974	1978	1979	1980	1981	1982	1983
1975	1.1						
1976	1.2						
1977	1.1						
1978	1.6						
1979	2.9	1.3					
1980	5.4	3.7	2.4				
1981	6.3	4.6	3.3	0.8			
1982	6.5	4.8	3.4	1.0	0.2		
1983	5.5	3.9	2.6	0.1	-0.7	-0.9	
1984	4.7	3.1	1.8	-0.6	-1.4	-1.6	-0.8

The last four columns of Table 2 are relevant to beneficiaries eligible under post-1977 law, which applies to retired-worker beneficiaries born after 1916, disabled beneficiaries who first became eligible after 1978, and auxiliary beneficiaries who became eligible on a post-1977 law primary account. Under the newer (and current) rules, the initial benefit is determined from a wage-indexed formula that applies for the year in which the beneficiary is initially entitled (62 for retirees, any age for disability beneficiaries). The initial benefit is then price-indexed for that and all subsequent years. Thus, a retired worker born in 1917 would benefit from CPI-W indexation beginning in 1979 and, as shown in the second column of the table, would have received annual benefits that, by 1982, were 4.8 percent higher than if benefits were indexed by the adjusted CPI-W. Assuming benefit receipt continued, the annual percentage overpayment would fall to 3.1 percent in 1984 and all subsequent years. If retirement continued through 1984 (1983 for the 1920-21 birth cohorts), post-1977 law retirees born during the years 1919-1922 received benefits *lower* than if the adjusted CPI-W were used to index retirement benefits. Retirees born after 1922 are unaffected, though disability and auxiliary beneficiaries born in a later year could be affected. Since auxiliary and survivor benefits are based on the benefit of the primary beneficiary/worker, the effect of the CPI bias, though not permanent, can extend for many years beyond the death of a primary beneficiary, as we demonstrate below.

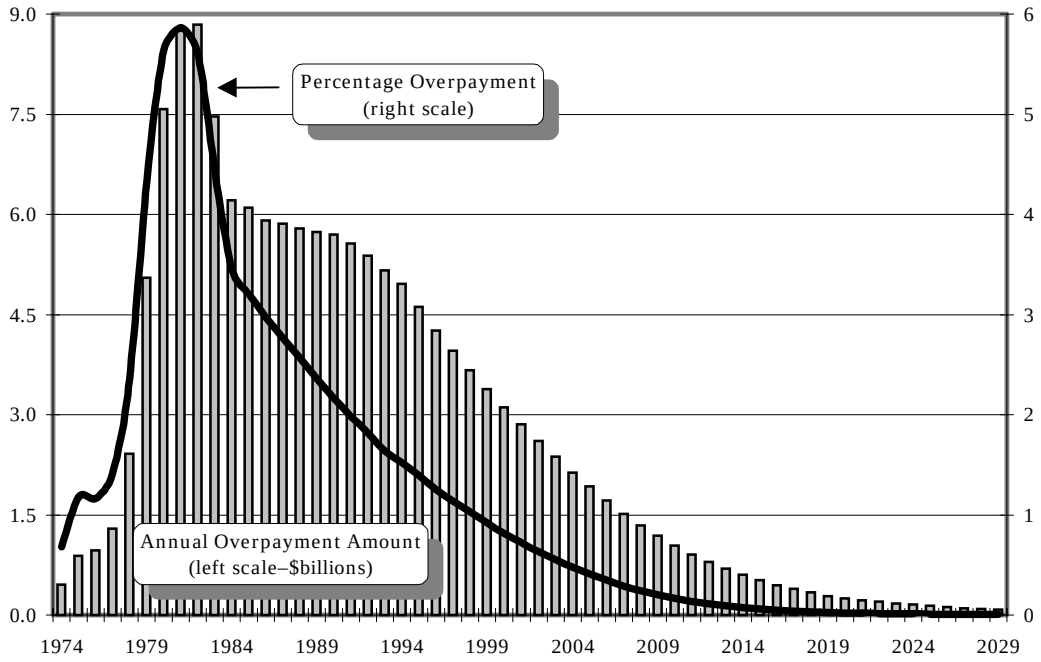
[3]Prior to 1974, discrepancies between the CPI-W and the CPI-U, X-1 were small, and, as noted in the previous section, Social Security benefits were not directly related to the CPI. Consequently, we need only to focus on changes in the CPI since 1974 in evaluating the effect of the mismeasurement of shelter costs.

3. EFFECTS ON THE BUDGET DEFICIT AND THE OASDI TRUST FUND

The CWHS-MBR provides the definitive means for estimating the effects of the CPI housing bias on both the federal budget deficit and the OASDI Trust Fund. This file contains a 1 percent sample of Social Security participants and all associated benefits paid in the period from 1975 to 1984 during which the CPI-W was subject to a housing bias. For computational purposes, we drew a twenty-percent random sample from the 2.7 million records on the CWHS file that we then matched with MBR records, yielding a sample of 536 thousand. For each of these observations, we estimated complete lifetime contributions and benefits (see Duggan and Gillingham (1999) for details). We then reconstructed lifetime benefits as they would have been paid if the price index used to compute benefits had been our adjusted CPI-W rather than the official CPI-W during the period 1975-84.[4] This data set allows our results to carefully reflect that, under current law, post-entitlement indexation errors have a temporary effect on benefits, lasting for as long as a given generation of primary and associated beneficiaries survives.

Figure 2 presents our estimates of annual benefit overpayments due to the housing bias in the CPI. The overpayments grow in absolute amount until 1983, and as a percentage of total payments until 1982. Since the bias was negative in 1983 and 1984, and corrected thereafter, the overpayments decay, both in level and as a percentage of payments, after these peaks.

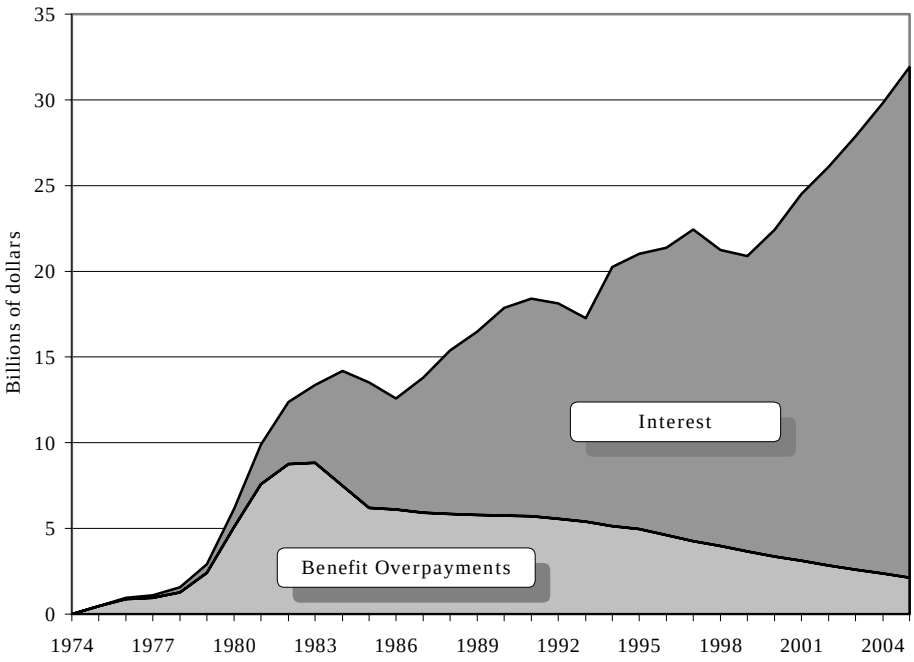
FIGURE 2. OVERPAYMENT OF OASDI BENEFITS, 1975–2030



[4] We can estimate the absolute size of the effect of the housing bias without imputing contributions and benefits for workers whose initial benefits were not effected by the housing bias. However, the more complete representation of benefits provided here places the effect of the housing bias in perspective relative to total benefits, rather than only those benefits actually affected by the bias.

Because the overpayments add to the budget deficit and reduce the OASDI Trust Fund, it is also important to measure the interest costs they imply. Figure 3 displays both the direct benefit overpayments and the associated interest costs. As the direct overpayments have decreased, the delayed interest effects have become the major cost to both the trust fund and the budget deficit. The total cost has grown to an estimated $20.9 billion in 1999, and should grow to more than $30 billion by the year 2005. Appendix Table A1 contains our annual estimates of total overpayments.

FIGURE 3. OASDI BENEFIT OVERPAYMENT AND INTEREST COSTS, 1975–2005

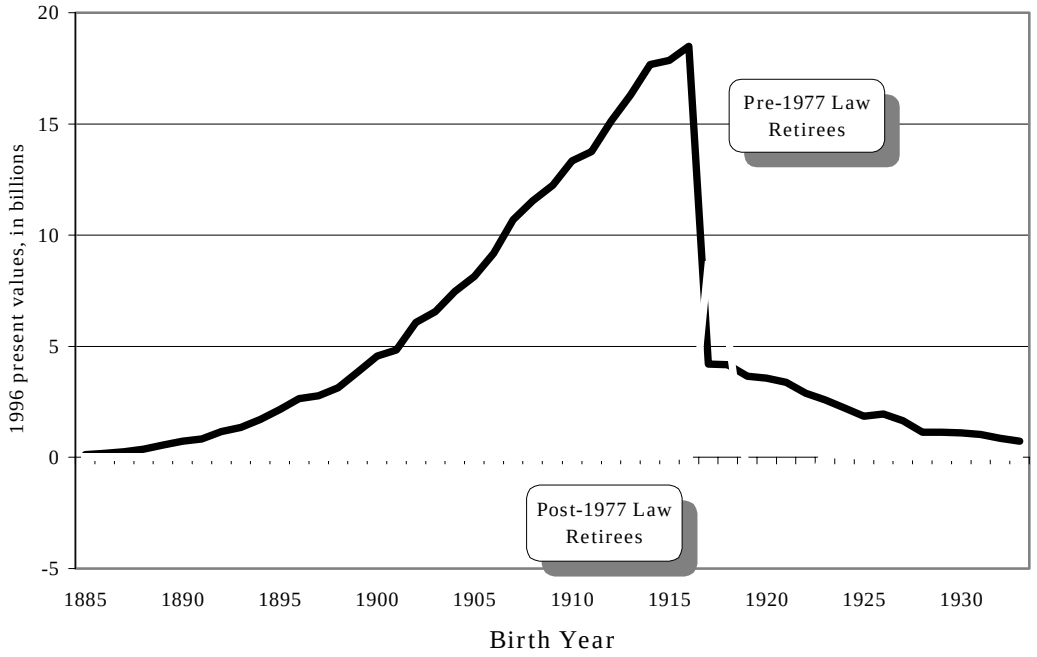


4. EFFECTS OF THE CPI BIAS ON INDIVIDUAL COHORTS

It is also interesting to ask which birth cohorts have benefited most from the housing bias. From our earlier discussion, we expect the answer to be beneficiaries eligible under pre-1977 law. Under that regime, the initial benefit formula was price-indexed and all beneficiaries would experience the full cumulative effect of the bias up to the time benefits terminate. Figure 4 presents our estimates by birth cohort of the present values of *total* benefit overpayments, discounted to 1996 with the interest rate paid or projected to be paid on the OASDI Trust Fund. Estimates are shown separately for pre-1977 and post-1977 Social Security law. Under the former, overpayments grow as the birth year approaches 1916, reaching a peak of $18.6 billion. Following the 1916 cohort, overpayments under pre-1977 law relate only to disability and auxiliary beneficiaries and therefore drop sharply beginning with the 1917 cohort. Nevertheless, all beneficiaries entitled under pre-1977 law who received benefits after 1974 were overpaid as a result of the CPI error.

The pattern of overpayments differs under post-1977 law. The 1917 and 1918 cohorts were overpaid but the next four cohorts received lifetime benefits lower than if the CPI-U, X1 had been used for indexation.[5] The first three underpaid cohorts belong to the group referred to as "notch babies" (birth years 1917-1921) because this group retired under rules governing a transition to post-1977 law that many have argued resulted in their unfair treatment. The underpayments identified here, however, are unrelated to the transition rules. In total, about 40 percent of post-1977 law beneficiaries affected by the CPI error were overpaid and 60 percent were underpaid. Small overpayments arise for auxiliary beneficiaries born after 1922 but soon go to zero.

FIGURE 4. PRESENT VALUE OF OASDI OVERPAYMENTS BY BIRTH COHORT PRE- AND POST-1977 SOCIAL SECURITY LAW



To set these results in a broader context, Duggan, Gillingham, and Greenlees (1993) found that all early participants (birth years 1895-1923) in the Social Security program received very high real rates of return on their contributions. For many pre-1977 law beneficiaries, those returns were higher than was intended due, in part, to the overindexation problem, which resulted in benefits growing faster than the cost of living. Figure 4 shows that, overindexation notwithstanding, returns were too high for pre-1977 law beneficiaries because the cost of living measure was biased upward during the relevant years. The payment error changes sign for early retirees under post-1977 law, including most of the notch babies. However, this underpayment only partially offsets the favorable treatment accorded these cohorts. In another

[5] Duggan, Gillingham, and Greenlees (1993) found that early participants (birth years 1895-1923) in the Social Security program received very high real rates of return on their contributions. For many pre-1977 law beneficiaries, those returns were higher than was intended due, in part, to the overindexation problem, which resulted in benefits growing faster than the cost of living.

study, Duggan, Gillingham, and Greenlees (1996) estimated that the notch babies as a group could expect to receive about $500 billion (1988 dollars) more in lifetime benefits than if they were paid the same rate of return on their contributions that the Social Security system earns on its invested funds. More importantly, they received higher net benefits than later cohorts and these relative overpayments swamp the effect of the CPI error.

FIGURE 5. MEAN PRESENT VALUE OF OVERPAYMENTS BY BIRTH YEAR RETIRED WORKER BENEFICIARIES

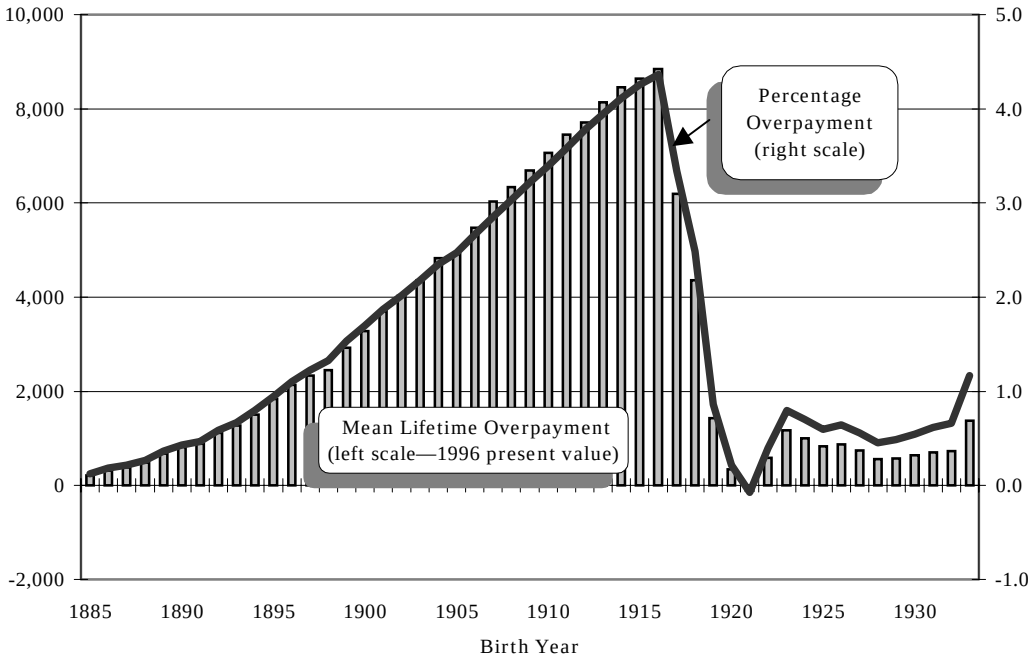


Figure 5 shows the pattern of *mean* overpayments per beneficiary (again expressed in present values as of 1996) for pre- and post-1977 law combined. The mean overpayment rises to 4.4 percent of benefits for the 1916 cohort before declining somewhat for the 1917-18 cohorts and then falling considerably beginning with the 1919 birth cohort. The overpayments persist well into the future, as seen in Appendix Table A2, accruing primarily to beneficiaries eligible under pre-1977 Social Security law when the related problem of "overindexation" prevailed. Summing these values over cohorts, the total effect of the housing error on the OASDI trust fund and, by implication, the federal debt, is just over $260 billion, discounted to 1996.

5. CONCLUSION

The purpose of this paper has been to provide a definitive estimate of the effects of an important, known bias in the housing component of the Consumer Price Index on the budget deficit and the Social Security trust fund. The estimated overpayment in 1999 is $3.7 billion, and, taking interest into account, the total drain on the OASDI Trust Fund and the budget is $20.9 billion. The 1996 present value of the entire effect of the error is $260.1 billion, a very

large amount, but relatively small in comparison with a gross federal debt at the end of fiscal year 1996 of over $5.2 trillion.[6]

The CPI bias resulted in substantial lifetime overpayments of OASDI benefits for cohorts eligible under pre-1977 Social Security law who received benefits after 1974. The overpayments will continue for many years into the future. The effect of the bias under post-1977 law was quite different. Overpayments resulted for retired and disabled workers (and their associated auxiliary beneficiaries) who became eligible during 1979-80. However, the CPI bias resulted in underpayments for beneficiaries who became eligible during 1981-1984.

[6] Throughout this paper, we have ignored indexation by the CPI in the federal budget other than for Social Security benefits. The individual income tax system has been indexed by the CPI-U since 1985, but this postdates the housing bias discussed here. The other significant use of CPI indexation is for federal civilian and military retirement benefits. This category totaled $72 billion in 1999, compared to $362 billion for Social Security benefits.

APPENDIX TABLE A1: OASDI BENEFIT OVERPAYMENTS (BILLIONS OF CURRENT$), BY YEAR

Year	Baseline Benefits	Benefit Overpayment	Percent Overpaid	Interest	Total Overpayment	Cumulative Overpayment
1975	67	0.5	0.7	0.0	0.5	0
1976	76	0.9	1.2	0.1	1.0	1
1977	84	1.0	1.2	0.1	1.1	3
1978	93	1.3	1.4	0.3	1.6	4
1979	105	2.4	2.4	0.5	2.9	7
1980	122	5.1	4.3	1.1	6.1	13
1981	142	7.6	5.6	2.3	9.9	23
1982	158	8.8	5.9	3.6	12.4	35
1983	167	8.8	5.6	4.5	13.4	49
1984	176	7.5	4.4	6.7	14.2	63
1985	186	6.2	3.5	7.3	13.5	76
1986	196	6.1	3.2	6.5	12.6	89
1987	205	5.9	3.0	7.9	13.8	103
1988	218	5.9	2.8	9.5	15.4	118
1989	232	5.8	2.6	10.7	16.5	135
1990	249	5.7	2.4	12.1	17.9	153
1991	269	5.7	2.2	12.7	18.4	171
1992	286	5.6	2.0	12.6	18.1	189
1993	301	5.4	1.8	11.9	17.3	206
1994	320	5.2	1.6	15.1	20.3	227
1995	332	5.0	1.5	16.1	21.0	248
1996	337	4.6	1.4	16.8	21.4	269
1997	345	4.3	1.3	18.2	22.5	292
1998	353	4.0	1.1	17.3	21.2	313
1999	361	3.7	1.0	17.2	20.9	334
2000	373	3.4	0.9	19.0	22.4	356
2001	385	3.1	0.8	21.4	24.5	381
2002	398	2.8	0.7	23.3	26.1	407
2003	414	2.6	0.6	25.3	27.9	435
2004	432	2.4	0.6	27.4	29.8	464
2005	452	2.1	0.5	29.8	31.9	496
2006	475	1.9	0.4	32.3	34.3	531
2007	500	1.7	0.3	34.0	35.7	566
2008	528	1.5	0.3	36.3	37.8	604
2009	559	1.3	0.2	38.7	40.0	644
2010	592	1.2	0.2	41.3	42.4	687
2011	629	1.0	0.2	44.0	45.0	732
2012	671	0.9	0.1	46.8	47.7	779
2013	717	0.8	0.1	49.9	50.7	830
2014	765	0.7	0.1	53.1	53.8	884
2015	816	0.6	0.1	56.6	57.2	941
2016	871	0.5	0.1	60.2	60.8	1,002
2017	929	0.4	0.0	64.1	64.6	1,066
2018	998	0.4	0.0	68.2	68.6	1,135
2019	1,074	0.3	0.0	72.6	73.0	1,208
2020	1,156	0.3	0.0	77.3	77.6	1,286
2021	1,244	0.2	0.0	82.3	82.5	1,368
2022	1,335	0.2	0.0	87.6	87.8	1,456
2023	1,434	0.2	0.0	93.2	93.4	1,549
2024	1,535	0.2	0.0	99.1	99.3	1,648
2025	1,639	0.2	0.0	105.5	105.6	1,754
2026	1,747	0.1	0.0	112.3	112.4	1,867
2027	1,857	0.1	0.0	119.4	119.6	1,986
2028	1,967	0.1	0.0	127.1	127.2	2,113
2029	2,076	0.1	0.0	135.2	135.3	2,249
2030	2,182	0.1	0.0	143.9	144.0	2,393

APPENDIX TABLE A2: OASDI BENEFIT OVERPAYMENTS BY BIRTH COHORT (PRESENT VALUES AS OF 1996)

	Mean Benefits			Total Benefits ($billions)	
Birth Cohort	Baseline Benefits	Overpayment	Percent Overpaid	Baseline Benefits	Overpayment
1885	171,210	216	0.1	100.1	0.1
1886	174,931	317	0.2	102.2	0.2
1887	179,923	385	0.2	117.9	0.3
1888	180,733	487	0.3	138.5	0.4
1889	187,350	683	0.4	151.3	0.6
1890	189,788	817	0.4	166.1	0.7
1891	189,177	885	0.5	178.0	0.8
1892	190,416	1,118	0.6	198.2	1.2
1893	189,572	1,271	0.7	203.5	1.4
1894	189,269	1,509	0.8	215.7	1.7
1895	194,402	1,836	0.9	228.4	2.2
1896	194,568	2,139	1.1	241.1	2.7
1897	190,892	2,333	1.2	228.0	2.8
1898	184,479	2,448	1.3	235.2	3.1
1899	191,355	2,931	1.5	251.7	3.9
1900	193,146	3,278	1.7	268.9	4.6
1901	198,129	3,704	1.9	259.0	4.8
1902	199,485	4,023	2.0	300.6	6.1
1903	199,851	4,361	2.2	300.1	6.5
1904	205,483	4,824	2.3	318.2	7.5
1905	199,070	4,923	2.5	329.5	8.1
1906	205,435	5,475	2.7	345.0	9.2
1907	211,036	6,024	2.9	374.3	10.7
1908	209,035	6,340	3.0	380.2	11.5
1909	207,818	6,695	3.2	380.5	12.3
1910	208,085	7,065	3.4	393.2	13.4
1911	207,855	7,455	3.6	384.0	13.8
1912	204,065	7,715	3.8	400.1	15.1
1913	206,280	8,137	3.9	413.8	16.3
1914	205,619	8,455	4.1	430.3	17.7
1915	203,109	8,647	4.3	420.2	17.9
1916	202,646	8,851	4.4	424.7	18.6
1917	185,459	6,194	3.3	384.0	12.8
1918	175,470	4,363	2.5	386.5	9.6
1919	165,797	1,430	0.9	349.3	3.0
1920	156,429	335	0.2	361.0	0.8
1921	157,794	(118)	(0.1)	375.4	(0.3)
1922	148,417	588	0.4	342.7	1.4
1923	147,206	1,172	0.8	341.4	2.7
1924	141,909	994	0.7	332.6	2.3
1925	140,423	836	0.6	318.2	1.9
1926	135,411	867	0.6	307.4	2.0
1927	131,791	738	0.6	298.5	1.7
1928	124,907	565	0.5	260.7	1.2
1929	118,833	579	0.5	237.2	1.2
1930	117,878	639	0.5	206.2	1.1
1931	113,445	698	0.6	174.1	1.1
1932	109,645	725	0.7	138.0	0.9
1933	117,505	1,374	1.2	59.2	0.7
Total for all cohorts between 1875 and 1933				14,149.9	260.1

REFERENCES

Advisory Commission To Study the Consumer Price Index. *Toward a More Accurate Measure of the Cost of Living*. Final Report to the Senate Finance Committee, December 4, 1996.

Duggan, J. E. and R. Gillingham, "The Effect of Errors in the CPI on Social Security Finances," *Journal of Business and Economic Statistics*, Vol. 17, No. 2, April 1999, pp. 161-169.

Duggan, J. E., R. Gillingham, and J. S. Greenlees, "The Returns Paid to Early Social Security Cohorts," *Contemporary Policy Issues*, XI, No. 4, October 1993, pp. 1-13.

Duggan, J. E., R. Gillingham, and J. S. Greenlees, "Distributional Effects of Social Security: the Notch Issued Revisited," *Public Finance Quarterly*, July, 1996, pp. 349-370.

Gillingham, R. "Estimating the User Cost of Owner-occupied Housing," *Monthly Labor Review*, Vol. 103, No. 2, February 1980, pp. 31-35.

Gillingham, R. "Measuring the Cost of Shelter for Homeowners: Theoretical and Empirical Considerations," *The Review of Economics and Statistics*, Vol. 65, No. 2, May 1983, pp. 254-265.

Greenspan, A. Testimony to the Committee on Finance, U. S. Senate, March 13, 1995.

Senate Finance Committee. *Consumer Price Index*, Hearings before the Committee on Finance, U. S. Senate, S.HRG. 104-69, U. S. Government Printing Office, 1995.

APPENDIX: DATA AND METHODS

We estimate the effect of CPI bias using the 1992 Continuous Work History Sample (CWHS) and the Master Beneficiary Record (MBR) files from the Social Security Administration (SSA). The CWHS is an earnings history sample for 1% of participants in the Social Security system and is the central building block for our simulation model. It contains over 2.7 million records with actual earnings histories spanning the period 1951 through 1992. The MBR is a benefit history file that contains associated OASDI benefits for the CWHS records.

We first draw a 20 percent sample from the CWHS, yielding 536 thousand observations. We then merge this sample with the MBR, to determine the benefit histories of those workers who have already started to receive benefits. Out of the total sample, 59 thousand workers have reached entitlement age, but have insufficient work histories (less than 40 quarters of coverage) to qualify for benefits. Of the remaining 477 thousand, 168 thousand have already started to receive benefits. For these workers, the MBR contains actual benefit records through 1995, which include not only the amount of the benefit, but also whether it reflects entitlement to a spousal, survivor, dependent, or dual-entitlement benefit. Spousal, survivor and dependent benefits are assigned as *auxiliary* benefits to the worker on whose contributions the benefits are based. A dual-entitlement benefit—a spousal or survivor benefit that goes to a spouse with her or his own benefit entitlement—is recorded on the benefit history of the spouse/survivor. In addition to their own—including dual-entitlement—benefits, the 168 thousand workers matched to the MBR also account for 51 thousand auxiliary benefits. For those recipients who exhaust all their benefits—including auxiliary benefits—prior to 1996, we have a complete accounting of their participation. For those recipients who retire and begin to receive benefits prior to 1996, we need only impute subsequent own and auxiliary benefits to have a complete accounting of their participation. For the remaining 309 thousand still in the labor force, we must complete both their work and benefit profiles.

Imputations for current workers. We proceed in two steps to impute the data necessary to fill out the contribution and benefit histories of our sample of workers. First, for current workers, we first impute the uncompleted portion of their *relative* earnings profiles based on information contained in the profiles of older workers. We focus on relative earnings, since we have exogenous estimates of the future time path of average earnings from the Social Security Trustees. At the same time that we impute future relative earnings, we also impute whether the worker becomes disabled, the profile of any disability episode, and whether the worker lives long enough to qualify for old-age benefits. For those whom we impute to qualify for old-age benefits, we use the completed earnings and disability profiles to project an initial benefit based on current-law benefit formulas.

Earnings imputation procedures. We use an imputation procedure to impute earnings that is related to the hot-decking procedure used by the Census Bureau to impute unreported income values in the Current Population Survey (see Ford (1983) for a more complete discussion of hot-decking procedures). Hot-decking refers to a set of imputation procedures in which missing values in a sample are imputed by duplicating reported values. In our application, values that we impute are not missing from failure to report; rather they have not yet been observed. There are several reasons why we use the hot-deck procedure rather than more traditional model-based approach to impute wages. First, the CWHS is unique in the large quantity of information it contains on actual historical life-cycle earnings. Second, the CWHS contains few

variables related to earnings, severely limiting the prospects for economic modeling of earnings. Finally, even with few "wage-equation" variables, a model-based procedure would be either overly simplistic or operationally intractable. By maintaining the assumption that the relative earnings profiles of slightly older workers are good proxies for the future relative wage profiles of current workers, we avoid having to model myriad parameters of the joint distribution of future labor force status, wages, disability, mortality and retirement.

To implement the imputation procedure, we classify our sample into smaller groups to increase within-group homogeneity. Missing values for younger cohorts—relative wages, disability and mortality—are replaced by reported values of slightly older cohorts within each classification group. The underlying assumption is that, within each group, the joint distribution of future earnings, mortality, and disability for younger cohorts is the same as that for slightly older cohorts.

There are two basic ways to choose a "donor" record from within a worker's classification group. *Sequential* hot-deck procedures, used by the Census Bureau, impute for a missing value the last reported value in the classification group read by the computer. The principal advantage of sequential hot-decking is efficiency in data processing. Alternatively, reported values in the classification group can be selected randomly. The appeal of random selection is that it puts a known probability mechanism into the imputation process and reduces the chances of multiple uses of the same donor record. We use random selection.

Hot-Decking and the CWHS. The major classifying variables available in the CWHS are age, sex, race, disability status and earnings. We group observations by sex, race (white, black, other), and disability status (whether disability benefits were ever received), resulting in 12 groups. For each group we compute the ratio of an individual's earnings in a given year to the mean earnings of their age group for that year. We then rank the observations into age- and year-specific quartiles of relative earnings. The quartile rank then becomes another classification variable in our imputations. We also compute for each year the ratio of mean age-group earnings to economy-wide wages. The two ratios, rather than earnings levels, are imputed.

Implicit in our imputations is the following identity:

$$w_{iat} \equiv c_{at} + p_{iat},$$

where w_{iat} is the log of relative earnings of the i^{th} person of age a in year t relative to the economy-wide average, c_{at} is the log of average relative earnings of workers of age a in period t relative to the economy-wide average, and p_{iat} is the log of relative earnings of the i^{th} person of age a in year t relative to the average for her or his age group. In our application, changes in c and p are implicitly functions of the joint distributions of human capital and random covariation embedded in the actual profiles. We assume that unobserved sequences of future changes in these variables for members of younger cohorts follow the same joint distribution as the actual sequence of changes in these variables over the same ages for members of slightly older cohorts.

As noted, our imputation groupings are defined according to age, sex, race, disability status, and relative earnings quartile. In choosing an observation from which to impute, we also condition on whether an individual is working when we last observe them and whether they have some earnings over the preceding five years. We imputed iteratively in reverse order of age using as possible donors individuals who have similar recent work experience, who were born within six years of the "donee," and have the same relative cohort earnings ranking at the

age imputation begins. When the donee record is complete, it becomes part of the potential donor file for younger age groups. For example, consider a white nondisabled continuously working male age 60 in 1992 as the "donee." The donor file consists of white nondisabled males born between 1925 and 1931 who have worked to at least age 60, have some work experience between ages 55 and 59, and have the same relative age-group earnings ranking at age 60. From this file a donor record is selected randomly. From the selected donor record the donee record gets, from age 61, future disability status; relative earnings ratios; ratios of mean age-group earnings to economy-wide wages; age of death if the worker dies before receiving old-age benefits and, if not, the year in which old-age benefits commenced. We begin imputations with the disability groups, which form roughly 5% of our sample. The disabled form a natural classification group, and when their profiles are completed they become part of the donor files for under-65, nondisabled workers, thereby allowing the latter group the possibility of becoming disability beneficiaries.

Benefit imputation procedures. We use two slightly different procedures for imputing benefits, depending on whether or not a worker has received benefits prior to 1996. For those who have, the MBR contains actual benefit records through 1995, which include not only the amount of the benefit, but also whether it reflects entitlement to a dual-entitlement or other auxiliary benefit. For those who exhausted all these benefits prior to 1996, the MBR provides a complete benefit history; for those still receiving benefits in 1995, subsequent benefits must be imputed. For those workers who receive initial benefits in 1996 or after, we must first calculate an initial benefit based on the worker's completed earnings history and current rules for calculating benefits. For these workers we have no direct information on whether they are entitled to spousal, dual-entitlement, or survivor benefits. Consequently, our estimates of future benefits for these recipients will be underestimated. The underestimation is mitigated by the fact that it starts at zero and increases slowly over the next 20 years. The present discounted value of *future* benefits in our simulations is roughly 10% less than comparable SSA benefit projections over the years 1998 to 2040.

Once we have a starting point for the imputation of future benefits, we complete these benefit histories with imputations based on current indexation rules and the age at which a beneficiary stops receiving benefits. For primary beneficiaries, we impute an age of death based on a logit mortality model estimated on the CWHS-MBR data set. The logit model is estimated separately for workers who have received disability benefits. Within disability beneficiaries, separate equations are estimated by race (white and other) and sex that relate mortality hazard to year of birth, age and the age at which disability occurred. For the nondisabled workers, the sample is again partitioned by race (black, white and other) and sex. In addition, the worker's average income from age 40 to 60 (in constant dollars), the number of years in which income exceeded the maximum on which contributions are levied, and whether the worker had no income between 40 and 60 are included as independent variables. Finally, similar equations, but with the income of the spouse, are estimated for all female beneficiaries to impute age of death for auxiliary beneficiaries. The estimated logit equations are summarized in Table A.1.

We estimate mortality endogenously to allow for the differential mortality by past disability status and income, in addition to the variation by age, sex, race, and birth year accounted for in Census mortality tables. For each beneficiary, the relevant logit model was initially evaluated at the age of the beneficiary and the year in either 1992 (for current recipients) or the first year of benefit receipt (for future recipients). A uniform random variable was then compared to the logit estimate of mortality probability. If the cumulated probability of death was greater than the random draw, the beneficiary was assumed to die in that year;

if not, the procedure was repeated with both year and age increased by one. For spouses, we use logit models that exclude income as an explanatory variable to impute an age of death, and then apply a similar random imputation procedure. For child beneficiaries, we assume benefits are exhausted at age 19.

TABLE A.1: ESTIMATED PARAMETERS OF LOGIT MORTALITY MODEL (ASYMPTOTIC STANDARD ERRORS IN PARENTHESES)

Population	6. Intercept	7. Age	Birth Year	Income	Years at income cap	No income	Age of disability
				Nondisabled			
Males							
White	8.3363	-.0839	.0204	.0581	.0315	.3574	
	(.1371)	(.0014)	(.0015)	(.0015)	(.0018)	(.0982)	
Black	7.3900	-.0741	.0085	.0810	.0076	.2878	
	(.4633)	(.0054)	(.0052)	(.0245)	(.0075)	(.2399)	
Other	9.2257	-.0894	.0071	.0408	.0376	.5083	
	(.8793)	(.0099)	(.0107)	(.0519)	(.0165)	(.4552)	
Females							
White	9.2626	-.0901	.0176	.0980	.0091	.2652	
	(.1742)	(.0020)	(.0022)	(.0093)	(.0043)	(.0866)	
Black	6.6928	-.0695	.0004	.1857	-.0487	.8124	
	(.5520)	(.0066)	(.0069)	(.0268)	(.0183)	(.2478)	
Other	8.1366	-.0837	.0152	.1898	-.0331	1.3915	
	(1.3366)	(.0159)	(.0184)	(.0654)	(.0350)	(.5286)	
				Disabled			
Males							
White	5.9817	-.0469	.0152				-.0074
	(.7006)	(.0078)	(.0078)				(.0042)
Black/Other	3.8494	-.0242	.0018				.0042
	(1.2949)	(.0146)	(.0127)				(.0112)
Females							
White	6.5893	-.0481	.0150				-.0058
	(.7349)	(.0072)	(.0068)				(.0080)
Black/Other	8.4615	-.0728	-.0087				-.0046
	(2.6282)	(.0287)	(.0288)				(.0209)

Note:
Income: Logarithm of average income (in constant dollars) between ages 40 and 60.
Years at cap: Number of years in which worker's income exceeded maximum on which contributions levied.
No income: Worker had no income from age 40 to 60.

APPENDIX REFERENCE

Ford, B. L. (1983), “An Overview of Hot-Deck Procedures,” in *Incomplete Data in Sample Surveys* (Vol. 2), eds. W. G. Madow, I. Olkin, and D. B. Rubin, New York: Academic Press, 185-207