The Response of Household Consumption to Income Tax Refunds

By NICHOLAS S. SOULELES*

A central implication of the life-cycle (or permanent-income) theory is that consumption should not respond to predictable fluctuations in income. Tests of this implication have yielded mixed results, especially on micro data (Angus Deaton, 1992; Martin Browning and Annamaria Lusardi, 1996). In large part this might be due to the difficulties of isolating the predictable component of income at the micro level. Most tests proceed by instrumenting for income, but since the available instruments are typically poor, such tests might be prejudiced against finding significant excess sensitivity of consumption to income (John Shea, 1995).[1] Also, it is not clear how closely the resulting econometric predictions of income coincide with agents' actual expectations of income.

To avoid these difficulties this paper examines the response of household consumption to a particular type of income that is both predictable and transitory—income tax refunds. Since a refund depends on events in the previous calendar year, it is predictable income as regards consumption in the year of its receipt. Consequently, under the life-cycle theory consumption should not increase on receipt of a refund.[2] In addition to testing the canonical model of consumption, this paper provides estimates interpretable as the marginal propensity to consume (MPC) out of refunds. Since federal tax refunds now amount to over $80 billion per year (averaging well over $1,000 per refund), these estimates are of interest in themselves. More generally they bear on the impact of even preannounced and temporary changes in fiscal policy.

The paper begins by surveying related studies in Section I. Section II describes the data, the Consumer Expenditure Survey (CEX), which of the leading U.S. micro data sets has the most comprehensive coverage of expenditure. The empirical specification is set out in Section III. Section IV reports the results, and Section V concludes.

I. Related Studies

Most studies of the response of consumption to fiscal policy have used aggregate, time-series data. Among these two employ the Euler-equation framework of Robert E. Hall (1978). First, James M. Poterba (1988) found a large increase in nondurable consumption in response to the 1975 tax rebate, 18 to 25 percent of the rebate within a month.[3] Second, Wilcox (1990)

* Department of Finance, The Wharton School, University of Pennsylvania, 2300 Steinberg Hall-Dietrich Hall, Philadelphia, PA 19104. I thank Jason Abrevaya, Joshua Angrist, Orazio Attanasio, Olivier Blanchard, Ricardo Caballero, Jerry Hausman, Jonathan Parker, Jack Porter, John Karl Scholz, Richard Thaler, David Wilcox, Steve Zeldes, two anonymous referees, and participants at various seminars for many valuable discussions. Special thanks go to Steve Pischke, who has been extraordinarily generous with his help throughout this project; and to the staff of the U.S. Bureau of Labor Statistics, for authoritative and patient help with the data. Of course, none of these people is responsible for any errors herein. An earlier version of this paper was part of my Ph.D. dissertation at MIT. Financial support from the National Science Foundation is gratefully acknowledged.

[1] Poor instruments can also lead to poor inferences in small samples (John Bound et al., 1995).

[2] This argument as stated makes a relatively strong—though quite standard—assumption about the rationality of people's expectations. The same conclusion can be reached, however, with relatively weaker assumptions. Consider a household that does not exactly know the amount of its refund until it completes its tax return in the spring. Any "surprise" at that time is unlikely to carry much information about the household's permanent income that was not already known by the end of the previous year. Therefore a refund is also a transitory kind of income, and so under the life-cycle theory there should still be little response in consumption on completing the return. Moreover, there should be no additional response on later receiving the refund. (In the period investigated in this paper, there was usually a 4–12 week delay between the filing of a return and refund receipt.)

[3] With only a few observations for the period of the rebate, the time-series specification used for this result was unavoidably constrained. Considering more generally some

investigated the time-series relationship between the disbursement of federal tax refunds and personal consumption expenditure. In his baseline specification he found an instantaneous consumption response on the order of 7.5 percent of refunds,[4] spread roughly equally across durables, nondurables, and services.

There are a number of advantages to employing the same framework at the micro level. First, aggregation bias can induce spurious excess sensitivity in aggregate data even when there is no such sensitivity in the underlying micro data (Orazio Attanasio and Guglielmo Weber, 1995). Second, there is a great amount of cross-sectional variation in refunds that time-series analyses neglect. Third, at the micro level one can investigate heterogeneity and nonlinearities in households' responses. For instance, one can contrast the response of households that are liquidity constrained and that of households that are not constrained.

There is an earlier tradition of studies at the micro level that examined windfalls. Most notably Ronald Bodkin (1959) found large MPCs from the national life insurance dividends paid to certain WWII veterans. (The estimated MPCs were 0.72 for nondurables, and 0.97 for total consumption, within a year.) By contrast Mordechai E. Kreinin (1961) found much smaller responses by Israelis to restitution payments from Germany (0.16 for nondurables, and 0.17 for total consumption). Part of the difference between these results might be due to misspecification resulting from working in levels of consumption, which working instead with Euler equations avoids.[5] There have not been many attempts since these studies to estimate MPCs in ways consistent with the modern life-cycle theory.

Since Hall's paper, most micro tests of the theory have used the Euler-equation framework applied to the response of consumption to instrumented total income. By contrast, since refunds are a predictable source of income, there is no need to instrument here. Coupled with the large cross-sectional variation in the magnitude of refunds, this should lead to a test that is more powerful than usual. Recently, a few other micro studies have similarly examined "experiments" like refund receipt in order to gain power. However refund recipients are a larger, more representative group than usually examined, and their response has aggregate implications for fiscal policy.[6,7]

Finally, the analysis of refunds also bears on two of the leading behavioral theories of saving. First, some leading expositions of "mental accounts" suggest that if a refund is large, it will be classified as a type of income to be saved. That is, the MPC out of a refund is posited to decrease in the size of the refund (Richard H. Thaler and George Loewenstein, 1989; Thaler, 1990).[8] Second, some theories of "self-control"

of the major changes in fiscal policy over the last few decades, Poterba found that consumption did not usually react on the announcement of the changes, as it should under the life-cycle theory, but only after their implementation. David W. Wilcox (1989) provided a similar result for preannounced increases in Social Security benefits.

[4] This is the estimated response on the very first day of refund receipt. The paper does not report the accumulated response over periods of greater length. The estimated instantaneous responses were somewhat sensitive to the particular instruments used for income.

[5] Bodkin was aware that his dividend variable might have been picking up the correlation of the dividend with omitted variables in turn correlated with permanent income. On adding likely such variables to Bodkin's original regression, Roger C. Bird and Bodkin (1965) found the MPCs decreased substantially, to 0.38 for nondurables, and 0.65 for total consumption. Also, note that the samples in the studies of windfalls are rather small (e.g., only 120 for Kreinin), and the recipients were rather special groups of people.

[6] Christina H. Paxson (1992, 1993) examined seasonal variability in consumption in Thailand; Shea (1995) examined the food consumption of union members in relation to changes in union contracts; and Souleles (1999) examined the effect on consumption of paying for college. More recently, two other micro papers have considered other fiscal experiments. Jonathan A. Parker (1999) looked at workers whose earnings hit the Social Security tax cap, and Souleles (1996) looked at the response to the Reagan tax cuts.

[7] Also at the micro level, Stephen K. McNees (1973) analyzed questionnaires of refund recipients. Forty-five percent of his sample said they spent their 1972 refund, 24 percent saved it, and 28 percent used it to pay off debt and bills. (The recipients were not asked how much of their refunds they applied to each use; nor were their spending, income, and assets recorded in the detail of the standard data sets on expenditure or assets.) Mathew D. Shapiro and Joel Slemrod (1995) analyzed a similar survey of the response to the change in withholding rates in 1993.

[8] It is interesting that some people (e.g., Kreinin, 1961) have cited a decreasing MPC as a possible explanation for

TABLE 1—FEDERAL INCOME TAX REFUNDS

	IRS data		CEX data
Fiscal year	Total refunds (billion $)	Average real refund (1982–1984 $)	Average real refund (1982–1984 $)
1979	43.0	763	845
1980	46.7	769	825
1981	53.0	816	899
1982	58.4	845	973
1983	62.0	829	902
1984	65.0	831	900
1985	68.9	846	912
1986	69.5	819	903
1987	69.6	767	851
1988	69.0	734	861
1989	71.7	711	771
1990	78.1	717	827
1979–1990			mean 874 median 561

Note: The fiscal year is the year before the refunds are received.
Source: Internal Revenue Service, *Statistics of Income: Individual Income Tax Returns,* Table 1.3, various years.

suggest that people deliberately overwithhold to force themselves to save, perhaps to save up enough for a durable.[9] Some implications of these theories for refunds will be investigated below.

II. The Data

A. *Income Tax Refunds*

As documented in Table 1, federal income tax refunds are commonly received and often large in magnitude. According to the Internal Revenue Service (IRS) about three-fourths of the 100 million or so tax filers each year receive refunds. Throughout the 1980's and early 1990's the average real refund fluctuated between about $700 and $850 (1982–1984 $). Table 2 shows the intrayear timing of refund disbursement, averaged across these years. In each year about 80 percent of refunds were mailed out in the three months of March, April, and May. There was, however, some variation in timing within these three months from year to year. For instance during the mid-1980's the disbursements came relatively late; most notably, in 1985 computer problems greatly delayed the processing of refunds in March, even though the IRS largely caught up by May.[10] This variation will be further discussed below.

B. *The Consumer Expenditure Survey*

The data are drawn from the CEX surveys for 1980 to 1991. In the CEX households are

the difference between the results of Bodkin and Kreinin: the Israeli restitution payments were relatively larger than the American insurance dividends.

[9] E.g., David D. Neumark (1995). This view is also frequently expressed in anecdotes and in the press.

[10] See Wilcox (1990). The table shows the distribution of the dollar value of refunds. While this is not the same as the distribution of the fraction of returns due refunds that are processed, the two distributions should be close to each other. Corroborating evidence at the level of individual returns comes from Slemrod et al. (1994), who look at the timing of the receipt of income tax returns by the IRS. Over 60 percent of returns due refunds are filed before April, mostly in February and March, with many of the rest filed in April. Applying the rule of thumb that in the years under consideration most returns were processed within 4–12 weeks, this evidence also suggests that most refund checks were received between March and May.

Taxpayers may credit their overpayment in a given fiscal year against their liability in the following year. It is unclear whether CEX respondents would include this credit when asked about refunds. (I guess most would not.) Doing so would tend to attenuate the results in the text. In any case, not many taxpayers do credit forward (e.g., only about 4 percent in 1991), so any resulting attenuation is unlikely to be significant.

TABLE 2—THE TIMING OF REFUND DISBURSEMENTS, 1980–1991

Month	Percent of annual disbursements
January	0.3
February	6.4
March	24.8
April	26.6
May	27.4
June	3.7
July	2.6
August	2.1
September	2.2
October	1.2
November	1.8
December	0.9

Notes: For each month and year, the fraction of the year's disbursements in that month was first calculated. These fractions were then averaged by month across the years 1980–1991. In months with a fifth Friday, half of the disbursements in the last week of the month were allocated to the following month.
Source: U.S. Department of the Treasury, *Daily Treasury Statements,* various issues, for individual federal income tax refunds.

interviewed four times, three months apart (though starting in different months for different households). In each interview the reference period for expenditure covers the three months before the interview. For each observation, a household quarter, individual expenditures were aggregated into real (1982–1984 $) consumption groups of increasing size, from food to total consumption. The U.S. Bureau of Labor Statistics (BLS) classification of nondurables (including services) includes many relatively durable and lumpy expenditures which are not as readily "smoothable" as the Euler equation assumes—for instance clothing. In response, following Lusardi (1996), a subset of nondurables, "strictly nondurables," has been created excluding such expenditures. The sample mean of strictly nondurables to total consumption is about 0.53.

The CEX asks about tax refunds twice, in a household's first and final interviews. Each time what is recorded is the value of federal tax refunds received by the household in the 12 months before the interview month. As with the other income variables in the CEX, there appears to be a nontrivial amount of underreporting: only about half of households report a refund, fewer than the approximately three-fourths of tax filers that receive refunds according to the IRS.[11] Table 1 compares over time the averages of the CEX refunds to the averages in the IRS data (both deflated). The CEX averages are greater, though reassuringly they display the same pattern across years as the IRS averages. The excess of reports of no refund, and the large means conditional on a nonzero report, might in large part reflect CEX respondents' rounding to zero (or forgetting) small refunds.[12]

The sample was selected in standard ways to improve the measurement of consumption. A household was dropped from the sample if there were multiple "consumer units" in the household, or if the household lived in student housing or the head of household was a farmer; a household quarter was dropped if the household lacked basic food expenditure for any month of the quarter, or if any food was received as pay in the quarter, and when the wording of the questions recording basic food expenditure changed in the following quarter. The sample was further restricted to households with heads aged 24–64.[13] The Appendix provides further details about the data.

III. Specification

A. *Excess Sensitivity*

The main equation that will be estimated is

$$C_{iy}^{II} - C_{iy}^{I} = \Sigma_s \beta_{0s} {}^* year_{is} + \boldsymbol{\beta}_1{}' \mathbf{X}_{iy} + \beta_2 {}^* refund_{iy} + u_{iy}. \quad (1)$$

The dependent variable is the change in a given household i's real consumption (in levels) between quarter I (January through March) and quarter II (April through June) of year y. An earlier version of this paper (available from the

[11] Further, one might expect household-level data to overstate refund receipt because a given household can file multiple tax returns.

[12] A histogram of the reported (nominal) refunds reveals that many households round their report to a multiple of $100, but otherwise the distribution seems reasonable.

[13] Students and those under 24 might still be dependent on their parents, and might not yet have a satisfactory understanding of how withholding works. Those over 64 generally do not have much income withheld and often have unusual and relatively poorly measured consumption, such as large health expenditures.

author) also examined consumption changes later in the year; the most salient of these results will be referred to below. Equation (1) nests a standard linearized Euler equation (when $\beta_2 = 0$), essentially that of Stephen P. Zeldes (1989) and Lusardi (1996). The vector $\mathbf{X}$ contains demographic variables (the age of the household head and changes in the number of adults and in the number of children) that help control for the most basic changes in household preferences. The year dummies help control for aggregate shocks and interest rates across time.[14]

The refund variable in the CEX has a reference period of 12 months, not 3 months like the consumption variables, which complicates the analysis.[15] In response the sample will be restricted to households whose refund reference period accommodates both the testing of excess sensitivity and, later, the estimation of a MPC. Specifically, the key regressor $refund_{iy}$ records the real (1982–1984 \$) value of refunds (in levels) that household i received in any 12-month reference period that covers the first two quarters of year y (when about 90 percent of refunds are received). For instance, for $y = 1990$ the reference periods must include all the months January 1990 through June 1990; accordingly, the allowable reference periods are January 1990 to December 1990, December 1989 to November 1990, ..., and July 1989 to June 1990. Because these reference periods do not extend into 1991, any refund recorded in $refund_{1990}$ is predetermined as regards the dependent variable C^{II}_{1990}–C^{I}_{1990}. More generally, the sample restriction ensures that the regressor $refund_{iy}$ is predetermined, and so under the life-cycle theory β_2 should be zero.

Equation (1) is estimated by ordinary least squares (OLS), with the standard errors corrected for heteroskedasticity. Unlike most other studies where the excess sensitivity regressor is current disposable income, here there is no need to instrument for the refund regressor because it is predetermined.[16] Of course this conclusion leans heavily on the assumption of rational expectations, in particular on the ability of households to calculate correctly their tax liabilities in advance. For comparison, equation (1) will also be estimated by two-stage least squares (2SLS). To improve precision the households not receiving refunds are added as a sort of "control group," with a refund regressor of zero. As discussed above, this group probably includes some households that did actually receive a refund but whose refund variable falsely reports no refund. In response the estimation is also undertaken excluding the control group.

B. *The Marginal Propensity to Consume*

In the spirit of John Y. Campbell and N. Gregory Mankiw (1990), the most natural alternative hypothesis is that households consume a fraction μ of their refund on receipt. μ is interpretable as the MPC out of refunds. One could estimate μ by replacing $\beta_2{*}refund_{iy}$ in equation (1) with $\mu{*}\Delta(refund)_{iy}$, where $\Delta(refund)_{iy} \equiv refund^{II}_{iy} - refund^{I}_{iy}$, the value of refunds received in quarter II of year y minus the value of refunds received in quarter I of year y. The CEX does not, however, record refunds at the quarterly frequency, and in any case most households do not receive refunds in both the first and second quarters of a given year.[17]

[14] Many other studies (e.g. Attanasio and Weber, 1995; Lusardi, 1996) also use time dummies instead of real after-tax interest rates. As a robustness check these rates were estimated for each household in the sample using the federal tax tables. However, introducing the estimated household-specific after-tax rates into the analysis had almost no effect on the results presented.

[15] Most excess sensitivity tests using the CEX or the Panel Study of Income Dynamics face difficulties due to various mismatches in the timing of the consumption and income variables, but few attempt corrections like those undertaken later in this paper. See, e.g., Shea (1995) or Lusardi (1996).

[16] Although not required for the excess sensitivity test, it would be interesting to consider the response to other types of income as well, in addition to refunds. However, such variables are not available in the CEX at the high frequency needed in this paper. In particular, it is not easy to contrast the response to refunds with that to final (April) tax payments. The tax variable is very often missing, resulting in a small sample. Also, the variable mixes together final payments and estimated taxes. This is quite unfortunate for the purposes of this paper, because estimated taxes are paid four times a year, so their timing is even harder to pin down than that of refunds; also, since they apply to the current year's tax liability, estimated taxes are not predetermined in advance of the time of their payment. I plan to further investigate taxes in future work.

[17] Taking the 12-month reference period literally, someone who received refunds both in spring of, say, 1990 and late in 1989 (but still within the reference period) might

Nevertheless, the regressor $refund_{iy}$ can be aligned with $\Delta(refund)_{iy}$ by using the distribution of aggregate refund disbursement, as follows.

To begin with, recall that most refunds are received in the spring, so any refund recorded in $refund_{iy}$ is relatively likely to have been received in quarter II of year y, and so $refund_{iy}$ is in fact already correlated with $\Delta(refund)_{iy}$. Of course some of the refunds in $refund_{iy}$ will have been received in other quarters (i.e., outside of April–June of year y).[18] Without correction these refunds will attenuate the estimate of the MPC; the estimated β_2 will be less than the true value of μ. To illustrate, suppose the sample consists of three people who each receive a \$1 refund; one person receives her refund in March, another receives it in April, and the third receives it in May. Suppose each person immediately consumes the entire refund, i.e., the true MPC $\mu = 1$. Then, in the dependent variable $C^{II} - C^{I}$ the \$1 consumed in March goes into the $-C^{I}$ term, and so partly offsets the \$2 spent in April and May that go into the C^{II} term. The estimated β_2 would therefore be $1/3$ (normalizing by the number of people), attenuated by a factor of $(2/3 - 1/3) = 1/3$ compared to the actual MPC.

Generalizing beyond this example, one can calculate an "attenuation factor" π from the distribution of aggregate refund disbursements: $\pi_{iy} = p_{iy}^{II} - p_{iy}^{I}$, where $p_{iy}^{I}(p_{iy}^{II})$ is the proportion of the refunds disbursed during i's refund reference period that were disbursed in particular in quarter I (quarter II) of year y. Multiplying the refund regressor by these factors should largely undo the attenuation.[19] The averages of these factors by year are listed in the notes to Table 5. The factors range from highs of over 0.5 in 1985 (due to the delay in disbursements due to computer problems) to under 0.2 in 1980–1981 and 1990–1991. The attenuation factors were also computed allowing for a two-week delay after the disbursement of refunds. This should conservatively accommodate any delay while refund checks were in the mail and before households cashed them, yielding a lower bound for the MPC. With this delay the factors $\pi_{(2\ weeks)}$ range from over 0.7 to around 0.4. The delay dampens the attenuation correction since fewer refunds are assumed to have been spent before April.[20]

IV. Results

A. *Excess Sensitivity*

Table 3 records the results of the main excess sensitivity tests for different consumption groups, with and without the control group (but not yet weighting by the attenuation factors). For expenditure on food, in the first column of results, the excess sensitivity coefficient β_2 on refunds is positive, though small at 0.016 and insignificant. The coefficient for strictly nondurables is larger, almost 0.03, and significant at the 5-percent level. Strictly nondurable consumption is, therefore, excessively sensitive to tax refunds—counter to the life-cycle theory. For nondurables the coefficient is slightly larger, though estimated with less precision.

report both refunds in $refund_{1990}$. However, such people are probably few in number, and according to the BLS staff, the CEX surveyor would probably discourage the reporting of the earlier refund.

[18] As Table 2 shows, most of these refunds would have come in March. But using March through May for the leading consumption period (instead of quarter II) would require that the lagged consumption period (instead of quarter I) extend back through December of the previous year, before the refunds were predetermined. This might generate false findings of excess sensitivity because there are often important innovations to income in December (e.g., raises, bonuses) that would be correlated with refunds in the following year.

[19] Of course this correction is only an approximation insofar as the timing of refund receipt varies across types of households. Slemrod et al. (1994) investigated this timing using IRS data. While there is some systematic variation in timing, the partial effect of a refund's magnitude on the filing date was surprisingly small (1.6 days earlier per \$1,000 of refund).

[20] A previous version of this paper corrected the attenuation using a different method, taking advantage of the fact that the refund question in the CEX is usually asked twice, nine months apart, so that the reference periods of the two questions overlap. For a few households comparing the two questions can further pin down the timing of refund receipt to after March of year y, thereby correcting most of the attenuation. Although the number of households for which this method is possible is small, the results were similar in magnitude to the main results reported in the text.

TABLE 3—EXCESS SENSITIVITY TESTS

	Food		Strictly nondurables		Nondurables		Total consumption	
Age	1.42**	1.05**	1.50	1.12	2.25	1.43	−0.38	2.07
	(0.63)	(0.48)	(1.03)	(0.77)	(1.66)	(1.21)	(4.87)	(3.43)
d(adult)	81.7**	117.0**	138.0**	164.4**	139.4*	145.7**	353.6*	323.9**
	(33.8)	(27.1)	(50.4)	(45.7)	(78.2)	(62.4)	(192.9)	(134.9)
d(kids)	69.8*	9.0	137.8**	51.9	137.2	14.3	428.8	116.2
	(37.7)	(31.8)	(51.8)	(45.5)	(155.8)	(103.6)	(299.2)	(207.6)
Refund	0.016	0.014	0.026**	0.024**	0.030	0.025	0.185**	0.184**
	(0.011)	(0.009)	(0.013)	(0.012)	(0.020)	(0.018)	(0.077)	(0.067)
Control group	no	yes	no	yes	no	yes	no	yes
Number of observations	4,121	7,622	4,121	7,622	4,081	7,525	4,081	7,525

Notes: The dependent variable is $C^{II} - C^{I}$, the change in consumption between the first and second quarters, for the indicated consumption groups. Coefficients on time dummies are not reported. Heteroskedasticity-corrected standard errors are in parentheses. The samples with the control group include the households not receiving refunds.

* Significantly different from 0 at the 10-percent level, **at the 5-percent level.

The coefficient for total consumption is significant and much greater, at over 0.18. That is, for each extra dollar of refund received, total consumption rises by 18 cents. Since total consumption is the sum of expenditures on nondurables and durables, it appears that most of the response to refunds comes in durables.

Adding the control group does usually decrease the standard errors a bit, as expected, without much changing the point estimates. This is consistent with the view that the underreporting of refunds reflects mostly smaller refunds. For brevity, subsequent tables will report only the results including the control group, and highlight the contrast between strictly nondurables and total consumption. Turning to the demographic variables, consumption increases with family size, as one would expect. The coefficient on age is generally less significant. Again for brevity, the coefficients for the demographic variables will not be reported hereafter since they turn out rather similar across specifications.

To see whether liquidity constraints are the source of the excess sensitivity, following Zeldes (1989) the sample is split to contrast the behavior of the households most likely to be constrained with that of the households least likely to be constrained. Tullio Jappelli (1990) has suggested that around 20 percent of U.S. households are liquidity constrained. To bracket this figure two splits are made along the distribution of liquid wealth (normalized by earnings). In a first, looser split the bottom 25 percent of the distribution is compared to the top 50 percent. The second, tighter split compares the bottom 15 percent to the top 25 percent.[21]

Table 4 begins with the loose split, in Panel A. The coefficients for food and strictly nondurables are quite a bit larger in magnitude for the constrained (in row [1]) than for the unconstrained (in row [2]), whose coefficients are close to zero. By contrast the response of total consumption is much larger for the unconstrained, and significant. These contrasts are even sharper under the tight split. For the constrained (row [3]) the response of strictly nondurables is now significant and large at over 0.10. [For nondurables the coefficient is 0.168 (0.046).] Even their response in food is large at about 0.06.[22]

Insofar as a relatively large response in

[21] Contrasting just the ends of the distribution, as in the "extreme split" of Zeldes, makes it more likely that what is being identified is the effect of liquidity constraints; in the middle of the distribution there would be many of both constrained and unconstrained households. The resulting sample sizes are smaller than one would expect on the basis of the original sample, because the wealth and income variables in the CEX, especially the wealth variables, are often missing.

[22] A previous version of this paper considered additional splits. For instance, splitting by income alone (taking the low income households as constrained) led to even sharper results than reported here, presumably in part because of the larger sample size (since income is less often missing than wealth). Alternatively, on bracketing age into ten-year intervals and interacting it with refunds, the young showed greater sensitivity to refunds, which is again consistent with liquidity constraints.

TABLE 4—LIQUIDITY CONSTRAINTS

Coefficient on refund	Food	Strictly nondurables	Total consumption
Panel A. Loose split			
[1] wealth/earnings < 0.025	0.032**	0.030	0.082
	(0.016)	(0.026)	(0.113)
Number of observations	1,038	1,038	1,030
[2] wealth/earnings > 0.08	−0.002	0.006	0.209**
	(0.006)	(0.009)	(0.107)
Number of observations	1,866	1,866	1,844
Panel B. Tight split			
[3] wealth/earnings < 0.01	0.058**	0.103**	0.011
	(0.019)	(0.029)	(0.134)
Number of observations	643	643	639
[4] wealth/earnings > 0.25	−0.006	0.000	0.261**
	(0.005)	(0.008)	(0.108)
Number of observations	937	937	925

Notes: The dependent variable is $C^{II} - C^{I}$, the change in consumption between the first and second quarters. The reported coefficients are those for the refund regressor. Coefficients on time dummies and demographic variables are not reported. Heteroskedasticity-corrected standard errors are in parentheses. Sample includes the control group, the households not receiving refunds.

* Significantly different from 0 at the 10-percent level, **at the 5-percent level.

nondurables by those with little wealth is indicative of liquidity constraints, these results constitute evidence that such constraints are important. On the other hand, the response in durables by those with substantial liquid assets requires a different explanation. While spending on durables is in part a form of saving, the timing of the spending remains puzzling: liquid households need not tie their durables purchases to the arrival of a refund check. Furthermore, an earlier version of this paper documented additional excess sensitivity in nondurables later in the year. This was related largely to the spending of refunds while on trips in the summer. This delayed sensitivity also requires a different explanation than liquidity constraints: the spending takes place well after households have received their refunds.

Turning to behavioral explanations, perhaps people are overwithholding to force themselves to save up enough for a durable. The problem with such a simple self-control story is that much of the spending on durables is by unconstrained households, who could have easily undone any forced saving by drawing down their liquid assets (or by taking out a loan, in particular against their refund).[23] To check whether the MPC is decreasing in the size of the refund, the refund squared was added to equation (1). The estimated coefficient on the quadratic term was negative for food, strictly nondurables, and nondurables. However, for total consumption it was significantly positive, counter to the simplest mental accounts story.[24] Since some expositions of mental accounts suggest the relative (not absolute) size of the refund is what matters, the quadratic term was also normalized by income. The coefficient for total consumption remained positive, but became insignificant. As a result one cannot come to any firm conclusion about the role of the relative size of the refund; the income variable might be too noisy.

B. *The Marginal Propensity to Consume*

To correct for the timing attenuation, the refund regressor in equation (1) is weighted by the attenuation factor π. The resulting coefficients

[23] See David Laibson (1997) for a clever model of self-control in the presence of liquid and illiquid assets.

[24] For total consumption the coefficient on *refund* is 0.038 (0.077), on *refund*^{2}*/10000* it is 0.072 (0.016). The canonical consumption model with uncertainty can also generate a declining MPC, but this model bears more directly on nondurables (Christopher D. Carroll and Miles S. Kimball, 1996). By contrast the mental accounts view appears to be applied to total consumption. On splitting the sample the quadratic term for total consumption is larger for the unconstrained than for the constrained.

TABLE 5—THE MARGINAL PROPENSITY TO CONSUME

	Food	Strictly nondurables	Total consumption	Durables	Nonvehicular durables
Refund*π	0.062**	0.093**	0.640**	0.537**	0.297**
	(0.031)	(0.037)	(0.224)	(0.225)	(0.137)
Number of observations	7,622	7,622	7,525	7,525	7,525
Refund*$\pi_{(2\ \text{weeks})}$	0.027	0.045**	0.344**	0.294**	0.128**
	(0.017)	(0.021)	(0.116)	(0.122)	(0.061)
Number of observations	7,622	7,622	7,525	7,525	7,525

Notes: The dependent variable is $C^{II} - C^{I}$, the change in consumption between the first and second quarters. The attenuation factors π and $\pi_{(2\ \text{week})}$ represent the probability the refund came in April–June minus the probability the refund came in January–March, as described in the text. $\pi_{(2\ \text{weeks})}$ allows for a two-week delay before the refund is received and spent. The average of the π's across households by year, for 1980–1991 respectively, are: 0.13, 0.16, 0.20, 0.36, 0.34, 0.55, 0.32, 0.28, 0.30, 0.21, 0.15, 0.16. For $\pi_{(2\ \text{weeks})}$ the respective averages are: 0.42, 0.45, 0.47, 0.56, 0.60, 0.73, 0.57, 0.57, 0.56, 0.55, 0.50, 0.50.

Durables equals total consumption minus strictly nondurables; nonvehicular durables equals durables minus vehicles. Coefficients on time dummies and demographic variables are not reported. Heteroskedasticity-corrected standard errors are in parentheses. Sample includes the control group, the households not receiving refunds.

* Significantly different from 0 at the 10-percent level, **at the 5-percent level.

can be interpreted as MPCs. As reported in the first panel of Table 5, the responses of food and strictly nondurables are significant and now rather large, at about 0.06 and 0.09. The response of total consumption is quite large at 0.64. That is, almost two-thirds of every extra dollar of refund is spent within a quarter. In the second panel, allowing for a two-week delay after disbursement results in smaller yet still substantial coefficients, about double the original coefficients in Table 3. The estimate for total consumption bounds from below the MPC at about 0.35. Even this amount is greater than most previous estimates of the impact of fiscal policy.[25,26]

The fourth column of Table 5 reports the results on estimating equation (1) for durables alone (defined as total consumption minus strictly nondurables).[27] As expected, the response of durables is approximately the difference between the responses of total consumption and strictly nondurables in the previous two columns. The final column looks at a subset of durables, nonvehicular durables (durables minus vehicles). Their response is significant and about half the size of the response of all durables. Thus there is significant spending of refunds on both vehicles and nonvehicular durables.[28]

For comparison with studies in which instrumentation is needed to make the excess sensitivity regressor predetermined, equation (1) was reestimated by 2SLS. Instrumenting for the refund regressor could also help control for, among other things, errors that households make in calculating their tax liability.[29] The

[25] The attenuation correction was also applied to the other analyses of the previous subsection, with essentially the same pattern of results as there.

[26] To compare to the aggregate data, the (deflated) consumption and refund variables in the CEX sample were aggregated across households year by year: $\Delta C_y \equiv \Sigma_i(C^{II}_{iy} - C^{I}_{iy})$ and $\Delta R_y \equiv \Sigma_i \pi_{iy} * refund_{iy}$. The resulting correlation corr(ΔC_y, ΔR_y) is 0.15 for nondurables, 0.53 for total consumption. The same correlations using the corresponding aggregate NIPA and Treasury data are similar though larger, 0.50 and 0.60, respectively. Instead aggregating the CEX sample only for households with a positive refund, corr(ΔC_y, ΔR_y) increases to 0.25 and 0.72, respectively. This increase suggests there might be some composition bias in using the usual aggregate data. Since there are only 12 observations in these time series, however, these results must be taken with qualification.

[27] This specification can still be interpreted as a Euler equation, in the absence of transactions costs as in Mankiw (1982). Since the refund is predetermined as of the lagged consumption period (quarter I), the MA(1) error structure Mankiw identifies is not an issue here.

[28] In simple probits and tobits of vehicle purchase, refunds are significant even controlling for income.

[29] Though, as argued above, any such errors are likely to represent only transitory innovations to income. Also, the summertime excess sensitivity discussed above came after the time of refund receipt, and so cannot be due to calculation errors.

TABLE 6—INSTRUMENTING FOR REFUNDS

	Food	Strictly nondurables	Total consumption	First-stage R^2
Refund*π	0.144	0.178	0.709	0.063
	(0.097)	(0.180)	(0.727)	
Number of observations	7,595	7,595	7,498	

Notes: The dependent variable is $C^{II} - C^{I}$, the change in consumption between the first and second quarters. Estimation is by 2SLS, with instruments $\mathbf{Z}$ = (age, age^2, education, number of dependents, tax status, mortgage status, income) for the refund regressor (weighted by the attenuation factors π as in the previous table). Coefficients on time dummies and demographic variables are not reported. Heteroskedasticity-corrected standard errors are in parentheses. Sample includes the control group, the households not receiving refunds.

* Significantly different from 0 at the 10-percent level, **at the 5-percent level.

instruments $\mathbf{Z}$ used here are essentially the variables used to explain refund receipt in a recent study by Joseph J. Cordes et al. (1995) using IRS data: $\mathbf{Z}$ = (age, age^2, education, number of dependents, tax status [single, joint, head of household], mortgage status [yes or no], and income).[30] The results appear in Table 6, again weighting by the attenuation factors (as in Table 5). The estimated coefficients increase somewhat in magnitude compared to those in Table 5, but are no longer significant. Instrumenting has greatly increased the standard errors, by factors of about 3 to 5. Of course, the instruments should be restricted to variables that do not themselves belong in the Euler equation. Unfortunately it is not clear which variables in $\mathbf{Z}$, if any, meet this restriction.[31]

V. Conclusion

This paper has found significant evidence of excess sensitivity in the response of households' consumption to their income tax refunds. Further, some of the sources of this sensitivity were identified. In particular, liquidity constraints appear to play an important role, because the nondurable consumption of constrained households increased at the time of refund receipt, far more than for unconstrained households. However, more than liquidity constraints are at play, because durables expenditure by the unconstrained also responded substantially; and the response of nondurables extended later into the year, after refund receipt. Furthermore, the response in durables by the unconstrained is not easily explained by standard models of durables or self-control, because liquid households could have brought their durables before receiving their refunds. There was also some evidence of a disproportionate response to larger refunds, counter to some views of mental accounts. Finally, instrumenting for refunds was found to decrease the power of the excess sensitivity test.

Having rejected the null hypothesis of the life-cycle theory, the paper estimated under a simple alternative hypothesis the marginal propensity to consume out of refunds. The response of total consumption was found to be at least 35 percent of refunds within a quarter, up to over 60 percent. Given the large aggregate value of tax refunds, these results imply rather substantial macroeconomic effects of refunds, and more generally of fiscal policy.

This analysis might be extended in a number of ways. First, given the importance of the response in durables, it would be valuable to model their discreteness explicitly (as in e.g., Attanasio [1994] or Janice C. Eberly [1994]). Second, other household characteristics, in addition to liquid wealth, that are informative about liquidity constraints might be used to better identify constrained households (Jappelli et al., 1998). Finally, one can similarly examine the response of consumption in other concrete

[30] The number of dependents and tax status were estimated by running the CEX households through the federal tax forms. Cordes et al. know directly whether a filer itemized his deduction; lacking this information, here itemizers are proxied by households having a mortgage. Cordes et al. have an additional variable, the fraction of total tax payments that is withheld, that is not available here; but it was insignificant in their specification.

[31] Instrumenting for the unweighted refund regressor led to a similar pattern of results; the resulting coefficients were somewhat larger than those in Table 3, but not significant. Instrumenting with only the mortgage indicator, for instance, instead of the entire vector $\mathbf{Z}$, also further increased the coefficient on total consumption, though decreased the coefficient on strictly nondurables, and again substantially increased the standard errors.

situations in which it is known in advance that income will change.

APPENDIX

In aggregating expenditures, if any component of a consumption group (food to total consumption) was topcoded or missing its cost, the whole group was set to missing. If any component of a group was missing its date or dated before the reference period, that group was dropped for all interviews for the household at issue. A large number of expenditures are dated in the month of the interview. Following the recommendation of the staff at the BLS, for consistency such expenditures were accrued to the following reference period. The major components of strictly nondurables are food; household operations, including monthly utilities and small-scale rentals (e.g., videos); apparel services and rentals; transportation fuel and services, including public transport; personal services; and entertainment services and high-frequency fees.

As regards refunds the survey asks, "During the past 12 months, did you (or any members of your CU [consumer unit]) receive any ... Federal income tax [refunds]? What was the total amount received by ALL CU members?" Topcoded values were not used. Refunds were deflated by the average of the monthly CPI for all items averaged over March, April, and May. Liquid wealth is (as usual in the CEX) the sum of amounts in checking and savings, when both are reported. All nominal variables were deflated to 1982–1984 $. A more detailed description of the data is available from the author.

REFERENCES

Attanasio, Orazio P. "Consumer Durables and Inertial Behavior: Estimation and Aggregation of (*S*, *s*) Rules." Working paper, Stanford University, October 1994.

Attanasio, Orazio P. and Weber, Guglielmo. "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey." *Journal of Political Economy*, December 1995, *103*(6), pp. 1121–57.

Bird, Roger C. and Bodkin, Ronald G. "The National Service Life-Insurance Dividend of 1950 and Consumption: A Further Test of the "Strict" Permanent-Income Hypothesis." *Journal of Political Economy*, October 1965, *73*(5), pp. 499–515.

Bodkin, Ronald. "Windfall Income and Consumption." *American Economic Review*, September 1959, *49*(4), pp. 602–14.

Bound, John; Jaeger, David and Baker, Regina M. "Problems with Instrumental Variables Estimation When the Correlation between the Instruments and the Endogenous Explanatory Variables Is Weak." *Journal of the American Statistical Association*, June 1995, *90*(430), pp. 443–50.

Browning, Martin and Lusardi, Annamaria. "Household Saving: Micro Theories and Micro Facts." *Journal of Economic Literature*, December 1996, *34*(4), pp. 1797–855.

Campbell, John Y. and Mankiw, N. Gregory. "Permanent Income, Current Income, and Consumption." *Journal of Business and Economics Statistics*, July 1990, *8*(3), pp. 265–79.

Carroll, Christopher D. and Kimball, Miles S. "On the Concavity of the Consumption Function." *Econometrica*, July 1996, *64*(4), pp. 981–92.

Cordes, Joseph J.; Galper, Harvey and Kirby, Sheila Nataraj. "Income Tax Overwithholding: Forced Saving or Transactions Cost?" Working paper, George Washington University, January 1995.

Deaton, Angus. *Understanding consumption.* Oxford: Oxford University Press, 1992.

Eberly, Janice C. "Adjustment of Consumers Durable Stocks: Evidence from Automobile Purchases." *Journal of Political Economy*, June 1994, *102*(3), pp. 403–36.

Hall, Robert E. "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence." *Journal of Political Economy*, December 1978, *86*(6), pp. 971–87.

Internal Revenue Service, Department of the Treasury. *Statistics of income: Individual income tax returns.* Washington, DC: U.S. Government Printing Office, various issues.

Jappelli, Tullio. "Who Is Credit Constrained in the U.S. Economy?" *Quarterly Journal of Economics*, February 1990, *105*(1), pp. 219–34.

Jappelli, Tullio; Pischke, Jörn-Steffen and Souleles, Nicholas S. "Testing for Liquidity Constraints in Euler Equations with Complementary Data Sources." *Review of Economics and Statistics*, May 1998, *80*(2), pp. 251–62.

Kreinin, Mordechai E. "Windfall Income and Consumption: Additional Evidence." *American Economic Review*, June 1961, *51*(3), pp. 388–90.

Laibson, David. "Golden Eggs and Hyperbolic Discounting." *Quarterly Journal of Economics*, May 1997, *112*(2), pp. 443–77.

Lusardi, Annamaria. "Permanent Income, Current Income, and Consumption: Evidence from Two Panel Data Sets." *Journal of Business and Economic Statistics*, January 1996, *14*(1), pp. 81–90.

Mankiw, N. Gregory. "Hall's Consumption Hypothesis and Durable Goods." *Journal of Monetary Economics*, November 1982, *10*(3), pp. 417–25.

McNees, Stephen K. "Tax Refunds and Consumer Spending." *New England Economic Review*, January–February 1973, pp. 3–11.

Neumark, David D. "Are Rising Earnings Profiles a Forced-Saving Mechanism?" *Economic Journal*, January 1995, *105*(428), pp. 95–106.

Parker, Jonathan A. "The Reaction of Household Consumption to Predictable Changes in Social Security Taxes." *American Economic Review*, September 1999, *89*(4), pp. 959–73.

Paxson, Christina H. "Using Weather Variability to Estimate the Response of Savings to Transitory Income in Thailand." *American Economic Review*, March 1992, *82*(1), pp. 15–33.

______. "Consumption and Income Seasonality in Thailand." *Journal of Political Economy*, February 1993, *101*(1), pp. 39–72.

Poterba, James M. "Are Consumers Forward Looking? Evidence from Fiscal Experiments." *American Economic Review*, May 1988 (*Papers and Proceedings*), *78*(2), pp. 413–18.

Shapiro, Mathew D. and Slemrod, Joel. "Consumer Response to the Timing of Income: Evidence from a Change in Tax Withholding." *American Economic Review*, March 1995, *85*(1), pp. 274–83.

Shea, John. "Union Contracts and the Life-Cycle/Permanent-Income Hypothesis." *American Economic Review*, March 1995, *85*(1), pp. 186–200.

Slemrod, Joel; Christian, Charles; London, Rebecca and Parker, Jonathan. "April 15 Syndrome." Working paper, University of Michigan, June 1994.

Souleles, Nicholas S. "Consumer Response to the Reagan Tax Cuts." Working paper, University of Pennsylvania, November 1996.

______. "College Tuition and Household Savings and Consumption." *Journal of Public Economics*, 1999 (forthcoming).

Thaler, Richard H. "Anomalies: Saving, Fungibility, and Mental Accounts." *Journal of Economic Perspectives*, Winter 1990, *4*(1), pp. 193–205.

Thaler, Richard H. and Loewenstein, George. "Intertemporal Choice." *Journal of Economic Perspectives*, Fall 1989, *3*(4), pp. 181–93.

U.S. Department of the Treasury. *Daily treasury statements*. Washington, DC: U.S. Government Printing Office, various issues.

Wilcox, David W. "Social Security Benefits, Consumption Expenditure, and the Life Cycle Hypothesis." *Journal of Political Economy*, April 1989, *97*(2), pp. 288–304.

______. "Income Tax Refunds and the Timing of Consumption Expenditure." Working paper, Federal Reserve Board of Governors, April 1990.

Zeldes, Stephen P. "Consumption and Liquidity Constraints: An Empirical Investigation." *Journal of Political Economy*, April 1989, *97*(2), pp. 305–46.