"April 15 Syndrome." Slemrod, Joel, Charles Christian, Rebecca London and Jonathan A. Parker. *Economic Inquiry* Vol. 35, No. 4 (1997): 695-709.
http://doi.org/10.1111/j.1465-7295.1997.tb01958.x

of doing their taxes. This model generates a wide distribution of filing times for returns which will get refunds and rationalizes procrastination. However, our framework fails to generate early filing of returns with taxes due, and we consider the issue again after our analysis of the data.

In section V, we take the predictions of our model to the data. We examine the various characteristics of refund and remittance returns and perform some reduced-form regressions of the determinants of return filing time. Most supportive of our interpretation of late filing for refunds, there is much less procrastination for returns which are filled out by a paid preparer, than for returns which are filed by the taxpayer. Further, we find that the complexity of forms delays filing, both for returns with tax due and for those with refunds. People with a higher marginal valuation of time, proxied by higher incomes, are more likely to file later, and the larger the refund value, the sooner the return is expected to be filed. Returns which are completed by professional tax preparers do not exhibit these last two characteristics. Finally, cross-year analysis demonstrates that filing late is habitual, suggesting that households have persistently different values of time or propensities to procrastinate. We conclude by speculating on the objectives and constraints which might generate early filing by individuals with taxes due and by considering the implications of our findings.

We believe that our investigation of the timing of tax return filing informs us about other examples of apparently non-optimal timing of economic behavior. Since people postpone the work of filling out tax forms despite real monetary costs, they may well postpone seeking a raise, comparison shopping, or switching supermarkets in response to price changes. Such behavior would generate "sticky behavior" leading to real rather than price adjustments in markets.

II. THE LEGAL FRAMEWORK

Individual tax returns for tax year 1988 were due on Monday, April 17, 1989.[2] For those who owed tax, the penalty for late *payment* was one-half a percent of the amount unpaid per month, not to exceed 25%. Further, interest, at roughly 11% per annum, was charged on the late taxes and on any penalty as it accrued.[3] For a return with taxes due that was *filed* late, there was an additional penalty of 5% of the amount due for each month (or fraction) the return was late, not to exceed 25%. For a return over 60 days late, a minimum penalty—the smaller of $100 or the amount of taxes due—became applicable. By filling out a simple form, the due date for the forms could be extended to August 15, but all taxes were still due on April 17. There was no penalty for filing for refunds late, but the forms must be filed within three years of the due date in order to claim the refund.

In order to avoid a penalty, estimated tax payments plus withholding had to equal the minimum of 90% of the current year's tax liability and 100% of the previous year's tax liability. The penalty for underpayment or late payment of estimated tax was calculated for three separate periods. For the first period (4/15/88 to 9/30/88), the penalty was 10% of the difference between the tax liability and the taxes paid, times the number of days in period 1, divided by 360. For periods two and three, (10/1/88 to 12/31/88 and 1/1/89 to 4/15/89, respectively) the rate was 11% per annum. If there was an underpayment in period one and a full payment in period two, a portion of the period-two payment was applied to the period-one payment and an underpayment was charged for period two.

Individuals choose how much to have withheld from their paychecks by filling out W-4 forms; these forms provide instructions for taxpayers on how they should be filled out. The only legal requirements and penalties associated with withholding are those associated with late or underpayment of taxes, as outlined above.

III. THE DATA AND THE DISTRIBUTION OF FILING TIMES

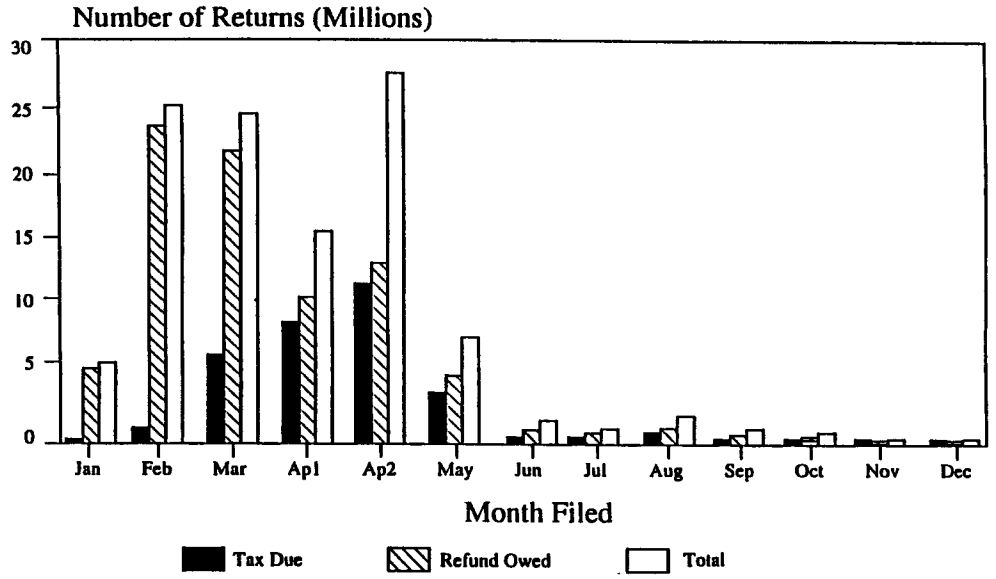
The Data

To evaluate the predictions of our subsequent model and to try to shed some light on why individuals pass up over $1 billion in in-

2. The deadline is April 15, unless that date falls on a weekend, in which case it is extended to the following Monday.

3. The interest rate was keyed to the short-term Federal rate for the previous quarter.

FIGURE 1
Number of Returns by Month Filed
All 1988 Returns



terest income by filing their tax forms when they do, we use the 1988 Internal Revenue Service Individual Model File. This data set is a stratified random sample of approximately 95,000 individual income tax returns filed during 1989 for tax year 1988 by U.S. citizens and residents. Each record has information on roughly 200 line items from the 1040 tax form and its supplementary schedules. To each record we have appended the date assigned by the IRS Service Center upon receipt of the return.

Filing Dates

Figure 1 shows the distribution of returns by month of receipt, with April split into halves, April 1 through 14 being denoted Ap1, and April 15 through 30 denoted Ap2. While 40% of all returns processed during the year are processed in April, it is clearly not the case that all returns with taxes due are filed at the last minute, nor is it true that all returns with refunds are filed well before the deadline. We also observe that the distribution of returns with refunds rises earlier than the distribution of returns with tax due, which is in accord with the economic incentives of the situation. Some 63% of tax returns with refunds are processed from January to March, while 19% of returns with taxes due are processed over the same period. When we looked only at returns with refunds or taxes due greater than \$500,[4] this difference was slightly more pronounced. Conversely, of returns with refunds less than \$500, 60% are processed from January to March, while of returns with taxes owed of less than \$500, 24% are processed then. Thus, whether the return has taxes owed or a refund seems to affect when it is filed. There also seems to be a size effect—people with larger refunds are likely to file earlier, while those with large amounts of tax due are likely to file nearer the deadline, although the shape of the distributions does not visibly change much when segregated by size of refund/remittance.

4. All dollar figures are reported in current, that is 1989, dollars.

We can also infer behavior from the distribution of returns filed after April 15. There is a rise in the number of returns processed in August, reflecting an "echo" effect similar to what we find in mid-April, for those people who file for automatic extensions and who face a second deadline of August 15. There is also an increase in the number of returns filed from the first to the second half of April. This reflects the April 17 deadline and possibly reflects IRS delays in assigning a date of receipt to returns received during the last-minute surge of filings in April.

There is one important caveat to our conclusions about the relationship between refund status and filing time, which one must also consider when interpreting our regression results in section V: taxpayers have significant control over whether they owe taxes or have taxes due. By changing their withholding during the previous calendar year or by making additional estimated tax payments, people can to a significant degree control the amount which they will receive as a refund or which they will owe. Why the great majority of taxpayers allow themselves to get into refund status, granting the government an interest-free loan during the tax year, is itself a fascinating question. For our purposes it implies that individuals may self-select into the categories which we interpret as explanatory variables. Because for some change in an exogenous variable (i.e., the elimination of a schedule), some individuals might switch refund categories, our estimates are not the deep structural parameters which determine the aggregate distribution of filing dates.

Ideally, we would correct for this bias by estimating a model which jointly determines the refund status and filing time, conditional on refund status. However, this requires either making strong assumptions about functional form or finding a variable which affects the decision as to which refund status to select into, but not the decision as to when to file. In light of the lack of a clear theory, we do not feel comfortable making either assumption. We refer the reader to two interesting preliminary empirical analyses of the withholding decision in Cordes et al. [1988a; 1988b]. Furthermore, in the regression analyses discussed below, those with a refund due are analyzed separately from those with tax due.

IV. CONCEPTUAL MODELS

In this section we sketch some simple models of when people fill out and mail in their tax returns, in the hope of matching the raw facts presented in the previous section. Our objective is to draw implications from these models which we can use to guide the empirical explorations that follow.

In the simplest model without uncertainty, consider a taxpayer who will receive a refund, i.e., taxes withheld plus estimated taxes paid exceed actual tax liability. Because it grants the government an interest-free loan, delaying filing costs the taxpayer foregone interest. Compared to filing at the last possible time T, the gain from filing before the deadline can be approximated by $iR(T-t)$, where i is the nominal interest rate, R is the refund amount, and t is the time of filing. If, alternatively, tax is due ($R<0$), it is optimal to file as late as possible; filing prematurely costs the taxpayer $iR(t-B)$, where B is the earliest possible filing date.

As Figure 1 makes clear however, the simple prediction of this model—a bimodal distribution of return filing dates at 0 and T—is not consistent with actual behavior. Instead, we observe a temporal distribution for returns which, while peaking at the deadline, is spread out between the earliest filing date and the deadline. We now examine some possible explanations for this pattern.

If, before investing time and money in the process, the taxpayer is uncertain about the sign and size of his balance due, it will generally be optimal to complete the tax forms as soon as possible, but in some cases postpone when the return is filed.[5] However, once the return is completed and it is known whether a refund is forthcoming, the decision as to when to file is trivial, and reduces to the problem above: file immediately if $R>0$, and wait until T if $R<0$. Considering uncertainty over the sign of the refund merely moves the puzzle from why people file when they do, to why they choose to acquire information when they do.

To us, a most plausible way to generate other than a bimodal distribution of filing

5. In fact many taxpayers do not need to complete the tax return in order to know whether a refund is forthcoming. Previous years provide an extremely good predictor, and quick calculations can often reveal into which camp a household will fall. Finally, as we pointed out in the previous section, returns with a large amount of taxes due still exhibit procrastination and substantial heterogeneity in filing times.

times is to allow the opportunity cost of completing a tax return to be a random variable, uncorrelated across time, but with substantial variation possible across individuals in mean and standard deviation. The idea is that each day taxpayers receive a random draw for their valuation of time that day, and decide whether to commit the time to do their taxes then or to postpone the task, hoping for a lower draw in the future. People do not know what future draws will be, nor can they use any draws prior to the current one, but they are aware of the distribution of their draws. This decision model does not allow for path dependence of utility and is perfectly rational and dynamically consistent.

Intuitively, this model corresponds to taxpayers making daily decisions about what to do with their leisure time, knowing that taxes have to be done by time T to avoid a penalty. Each day taxpayers consider other leisure options, and their current mood. We view these random values of leisure time as resulting from a combination of tastes and external options for leisure time activities. One is unlikely to fill out tax forms on an evening one has been invited out with friends, or on an evening when one is feeling particularly lazy and unfocused.

This model leads to a non-degenerate distribution of tax return completion times, and parallels job-search models in the labor literature. The solution is characterized by a reservation level of opportunity cost, call it z_t, such that the taxpayer who reaches time t with the tax forms uncompleted will choose to complete the forms if the draw of opportunity cost that day is lower than z_t; otherwise the taxpayer will postpone doing taxes for another day as in McCall [1970] and Lippman and McCall [1976]. If, upon completion of the tax return, the taxpayer is owed a refund, the return should be filed immediately; if tax is due, the return should not be remitted until the filing deadline.

The predictions of this model fit the relevant aspects of observed behavior. First, people do postpone filling out their tax forms, even when they can reasonably expect a refund, sometimes waiting until just before the deadline. When the deadline is approached they often will be filling out their forms in a high disutility-of-time period, having passed up lower disutility periods, thinking they would get lower draws still. If people receive high enough draws as they near the deadline, it may be optimal for them to take the penalty and file late, allowing them to fill out their forms at a more convenient time; filing late can be optimal behavior. Thus, this model generates a non-degenerate distribution for filing of returns with refunds due.

Further predictions of this model, which follow directly from the solution of the analogous model presented in Mortensen [1977] and extended in Engberg [1991], are as follows. An earlier expected time of filing is associated with a higher reservation level of opportunity cost relative to the expected value of opportunity cost, which would be caused by a higher refund, lower costs of filing, higher interest rates, or a lower subjective discount rate. Thus, our model predicts that people with lower valuations of time, which in section V we will represent by lower incomes and age exemptions, should file earlier, for a given value of refund. Secondly, on average those people with larger refunds should file earlier than those with smaller refunds, who should file sooner than those people with taxes due, ceteris paribus. We expect complexity of the return to increase the time involved in completing the forms, thus causing people to file later. Note that in the simple model discussed above an increase in filing cost affected only people on the margin between filing at the earliest possible time or at the last minute; by contrast, in this model complexity decreases everyone's probabilities of filing in early periods. People with higher effective discount rates should postpone filling out their tax forms longer.

Before leaving the discussion of the conceptual model, one point deserves further emphasis. The model outlined here can explain a distribution of filing times for returns filed with refunds. However, neither this model nor any model with rational behavior can easily explain why returns with tax balance due are filed before the deadline. It may be that some taxpayers are facing the penalty for underpayment which, as described in section II, depends in part on the filing date.[6] It also may

6. We are grateful to Margaret Reed for alerting us to this possibility.

be that early filers with a balance due are concerned that the return will be misplaced if not mailed immediately, are concerned that funds will not be available at a later date, or are extremely averse to being in debt. We delay further speculation on these points until after our regression analysis.

In his 1991 address to the American Economics Association, George Akerlof emphasized the potential importance of procrastination for economic behavior and set out a simple model of the phenomenon. It is interesting to contrast his model of procrastination with the model outlined above. In Akerlof's model there is a high relative (negative) weight placed on work today relative to work tomorrow—the agent is endowed with a high one-period discount rate. Thus people have a strong preference to postpone tasks until the next day. As the days pass this can lead to lengthy procrastination resulting from a series of small decisions. Akerlof's model implies that if it is optimal to postpone completing one's tax forms for one day at $t = 0$, it is optimal to postpone, day by day, until $T - 1$. If it is still optimal to postpone at $T - 1$, then it is optimal never to file one's taxes (barring criminal penalties, etc.).

The critical difference in the two models is their dynamic consistency. The model presented here is dynamically consistent, while the Akerlof model is dynamically inconsistent—a fully rational agent could see that any decision to postpone by just one day would lead to further procrastination and would not plan to postpone by just one day.[7] Second, our model rationalizes the observed distribution of tax filing times. If individuals were interviewed at the end of January and asked to assign probabilities as to when they would fill out and file their taxes, these answers would differ significantly depending on which is the correct model. The model outlined in this paper predicts that people understand that they are likely to postpone doing their taxes, and could give reasonable approximations to the true probabilities of their postponing their taxes until the week before April 15th. People whose behavior is governed by the Akerlof model would respond that they plan to do their taxes in the next week, while they would actually often postpone this task until the filing deadline is near.

7. The optimal dynamically consistent strategies to similar problems have since been analyzed in Laibson [1994].

V. EMPIRICAL RESULTS

Characteristics of Early and Late Filers

Table I presents information about the average tax return characteristics of taxpayers that filed during three different periods of 1989, classified by whether the returns had refunds or tax due. We name these taxpayers "early filers" if their forms were processed from January to March, "procrastinators" if their returns were processed in April, and "late filers" if their returns were processed after May 1. The table displays the means of several tax return items and also the percent of the forms filed with certain characteristics.

Several patterns are visible in Table I. Among those returns with a balance due, people who file earlier on average have lower amounts of tax due. The parallel prediction that early filers have larger refunds is not supported by this table. Second, as we hypothesized, complexity is apparently associated with late filing. The percentage of returns with interest and dividend income, supplemental income, farm income, capital gains income, itemizing, Keogh, self-employment income, estimated tax payments, married status, and using Form 1040 (rather than the simpler Form 1040A or Form 1040EZ) all increase as one moves from early to late filers. The percentage of returns which are 1040A and 1040EZ declines through time. Although not part of our set of hypotheses, we find that 40% of those early filers with taxes due are elderly while only 11% of tax returns are filed by the elderly. Most of these trends occur for all returns, for returns with refunds, and for returns with tax due.

The average income rises across the three periods, which may be an indication of the correlation between income and complexity, and/or that higher income individuals have a higher cost of time and are more likely to postpone filling out their forms. Note that for returns with taxes due, adjusted gross income is much higher for those who file in April, while for returns with refunds it increases concomitantly for late filers. This fits our theory. Returns with refunds have no monetary

TABLE I

Average Tax Return Characteristics by Time of Processing and Refund Status

	All Returns			Returns with Tax Due			Returns with Refunds		
	Jan–March	April	May–Dec	Jan–March	April	May–Dec	Jan–March	April	May–Dec
% of Total Returns	53.5	38.2	8.3	4.6	17.2	2.7	48.9	21.0	5.6
% with Refund	91.3	54.9	68.0	0	0	0	100	100	100
*Amount of Refund or Tax Due**	664	–600	460	1015	2468	3017	823	937	2098
Adj. Gross Income	20897	34090	48615	24018	43737	52306	20602	26156	46876
Wages and Salaries	18334	23910	29488	11951	26082	28194	18946	22124	30097
% Filing Schedule B	2.8	8.4	10.5	7.1	10.2	8.9	2.4	6.9	11.2
*Interest and Dividends***	8177	15000	32111	11143	17173	24459	7341	12237	34976
% Filing Schedule C	5.1	17.7	29.3	9.6	23.1	40.8	4.7	13.2	23.9
*Profit/Loss from Business***	2620	10407	14244	7915	14098	19367	1602	3918	10112
% Filing Schedule E	6.3	18.2	31.0	14.2	17.7	20.0	5.5	15.9	31.0
*Supplemental Income***	564	5338	6629	7645	22410	35482	–959	185	5228
% Filing Schedule F	1.8	2.1	2.3	9.1	2.5	2.7	1.1	1.8	2.1
*Farm Income***	3091	–2357	–9937	12798	130	–1450	–4342	–5118	–15122
% Filing Schedule D	4.6	12.9	17.5	14.2	17.7	20.0	3.7	8.9	16.3
*Capital Gains***	3700	15737	40545	7645	22410	35482	2272	4825	43471
% Married Filing Jointly	39.5	49.0	49.2	45.5	55.4	54.2	38.9	43.8	46.8
% Itemizer	21.8	36.0	43.7	18.3	39.3	47.1	22.2	33.2	42.1
% with Form 1040	52.7	77.3	85.2	83.3	86.4	90.5	49.7	69.8	82.7
% with Form 1040A	23.5	11.1	7.3	9.7	7.9	5.1	24.9	13.7	8.4
% with Form 1040EZ	23.8	11.6	7.5	7.0	5.7	4.4	25.4	16.5	8.9
% with IRA	4.6	12.9	17.5	6.2	8.1	6.0	4.5	7.6	6.4
% with Paid Tax Preparer	39.4	53.5	63.3	58.5	58.3	67.4	37.6	49.5	61.3
% with Estimated Tax Payments	6.0	16.5	22.7	24.3	22.4	21.6	4.3	11.7	23.2
% Elderly	10.8	14.2	10.4	38.8	18.4	8.4	8.1	10.7	11.4

*For all returns category, positive number indicates refund, negative number indicates balance due.

**Average over those who file particular schedule.

penalty associated with late filing; therefore, of people who have not filed early and receive low draws of opportunity cost near the deadline, those with refunds are more likely to procrastinate further. Further, some of the remittance returns processed in early April may have been completed in a low opportunity cost period early on, and then held to be mailed in at the deadline.

Table II reveals that these patterns also appear when the preparation status of the return is held constant. It also shows clearly that returns prepared by a professional are on average more complex and the filers have higher income.

Regression Analysis

Drawing conclusions based on these patterns is problematic because there are significant correlations among many of the categories on the left-hand side of the table. However, we can examine the partial associations of filing behavior with return characteristics through multiple regression analysis. The structure which our theory imposes on the data suggests we estimate a structural search model. However, we chose not to employ this procedure because of its difficulty in estimation and, in particular, the sensitivity of such models to errors-in-variables and omitted-variables bias. Wolpin [1987] and Engberg [1991] both demonstrate and comment on the complexity and sensitivity of this method.

As an alternative, we employ a Tobit regression. We eliminate from the sample all returns processed after April 30, in order to examine primarily those taxpayers who did not file late.[8] All our conclusions hold for this sample only, and do not take into account the phenomenon of late filers becoming procrastinators. However, our results do not change much if we set this cutoff date to, for example, December 31 or to April 22. We treat all returns processed on April 17th or later as if we did not know the true desired date of filing, but observed only that this date was constrained by the deadline.[9] The Tobit regression is weighted by the population weight of the individual.[10]

Table III presents the results of these regressions, both with and without a dummy variable for paid preparer status. With the exception of adjusted gross income and the presence of an Individual Retirement Account or Keogh plan, all the variables in Table III have estimated coefficients which are significantly different from zero at the 0.1% level.[11]

According to the results presented in Table III, filing later is associated with larger estimated payments and a larger balance due, and with using supplemental schedules (with the exception of Schedule F). The association with balance due provides some evidence of economic rationality and is consistent with model predictions. The magnitude of the coefficient is, however, small, suggesting that it takes an extra $4525 refund to speed up the filing of the return by one day. The positive coefficients on the supplemental schedules, which are interpreted as a dimension of complexity, are also consistent with model predictions. Filing earlier is associated with filing jointly, being 65 or older (also consistent with model predictions), and using either the Form 1040A or Form 1040EZ "short forms."

We refrain from drawing inferences from the coefficient on preparation mode (self-prepared or contracted out) because of taxpayer self-selection of this choice variable. The functional relationship between filing date and preparation mode may be further complicated by the potential endogeneity of preparation mode. It is possible that procrastinators resort to a paid preparer at the last minute in an attempt to file a timely return. In order to investigate the importance of these effects, we also offer reduced-form estimates of the model, excluding the preparer variable, in

8. The filing deadline was April 17. However, there is a variable lag between when the return was marked and when it is processed by the IRS.

9. Our results also did not change much when we replaced April 17 with April 10.

10. Population weights are calculated by the IRS by dividing the population count of returns in a sample stratum by the number of sample returns for that stratum. Strata are primarily income-based.

11. At the suggestion of a referee, we checked the validity of some of the structure imposed by the Tobit assumption. For the all-returns regression, we ran both a probit, in which the dependent variable denoted whether a return was censored, and an ordinary least squares regression on the non-censored population. The probit coefficients generally replicated the Tobit coefficients in sign and relative magnitudes, with the exception of *Paid Preparer*, which became small and positive. In the OLS regression only *Age* and *Keogh Plan* flipped sign.

TABLE II

Average Tax Return Characteristics by Time of Processing and Preparation Status

	Self–Prepared						Paid Preparer					
	Tax Due			Refund			Tax Due			Refund		
	Jan–March	April	May–Dec	Jan–March	April	May–Dec	Jan–March	April	May–Dec	Jan–March	April	May–Dec
% of Total Returns	1.9	7.2	0.9	30.5	10.6	2.2	2.7	10.1	1.8	18.4	10.4	3.5
Amount of Refund or Tax Due	455	987	1399	697	621	906	1412	1905	2250	1033	1259	2851
Adj. Gross Income	21840	33961	35938	17682	20894	25291	25561	50717	60230	25444	31519	60498
Wages and Salaries	14124	25835	25935	16598	18954	20310	10412	26257	29290	22839	25356	36275
% Filing Schedule B	4.4	6.2	4.1	1.2	3.1	4.4	9.0	13.1	11.3	4.4	10.7	15.6
*Interest and Dividends***	11797	9520	11207	6358	8211	17772	10914	19756	26762	7788	13558	38024
% Filing Schedule C	5.7	14.7	27.1	2.3	7.6	16.2	12.3	29.1	47.5	8.8	18.9	28.7
*Profit/Loss from Business***	5793	9653	14067	1468	2417	5206	8613	16803	20831	1658	4531	11854
% Filing Schedule E	6.8	10.3	13.4	2.3	7.3	13.9	20.6	28.7	39.6	10.7	24.7	41.8
*Supplemental Income***	4554	5930	11818	–1210	–1227	3527	6963	11133	9237	–869	610	5586
% Filing Schedule F	2.7	0.9	0.9	0.3	0.5	0.6	13.6	3.6	3.6	2.6	3.2	3.0
*Farm Income***	4999	–553	487	–3695	–3058	–6939	13912	255	–1675	–4455	–5466	–16201
% Filing Schedule D	6.5	10.9	10.1	1.9	4.8	8.6	19.7	22.6	24.7	6.7	13.1	21.1
*Capital Gains***	4073	8911	17021	1589	24.77	13742	8482	27053	39148	2596	5706	51069
% Married Filing Jointly	37.4	49.6	46.5	30.0	34.0	35.4	51.2	59.6	57.9	53.6	53.8	54.0
% Itemizer	15.2	33.3	38.2	14.6	23.8	27.8	20.5	43.6	51.3	34.6	42.8	51.2
% with Form 1040	65.8	70.7	77.0	30.3	48.2	61.8	95.8	97.7	97.0	82.0	91.9	96.0
% with Form 1040A	18.6	16.2	9.7	30.7	20.9	16.5	3.3	1.9	2.9	15.1	6.3	3.2
% with Form 1040EZ	15.6	13.1	13.3	39.0	30.9	21.7	0.9	0.4	0.1	2.9	1.8	0.8
% with IRA	5.6	6.1	6.5	3.6	6.0	5.6	6.6	9.5	5.8	5.9	8.0	7.0
% with Estimated Tax Payments	17.1	12.0	12.2	2.0	5.5	12.0	29.5	29.9	26.2	8.0	18.0	30.3
% Elderly	30.9	12.7	5.4	5.2	6.7	7.7	44.4	22.5	9.9	12.9	14.8	13.7

**Average over those who file particular schedule.

TABLE III
Regression Estimates
(Weighted Tobit)

Calendar Year 1988 Returns Filing During January–April 1989
Dependent Variable—Filing Date (Censored at 4/17/89), January 1 = 1, etc.

	w/PREP	Standard Error	w/o PREP	Standard Error
Intercept	83.3***	0.36	84.2***	0.30
Agi	0.0282	0.50	0.0251	0.050
Estpay	0.141***	0.034	0.144***	0.034
Baldue	0.221***	0.025	0.223***	0.025
Mfj	–3.60***	0.31	–3.62***	0.31
Age	–2.72***	0.43	–2.70***	0.43
Scha	3.24***	0.37	3.187***	0.37
Schb	5.96***	0.68	6.04***	0.68
Schc	12.1***	0.61	12.2***	0.61
Schd	7.46***	0.56	7.51***	0.56
Sche	10.9***	0.45	11.1***	0.45
Schf	–7.58***	0.97	–7.35***	0.96
Schse	9.73***	0.67	9.85***	0.67
IRA	0.383	0.56	0.307	0.56
Keo	0.222	1.91	0.150	1.91
F1040A	–12.6***	0.41	–13.1***	0.39
F1040EZ	–15.0***	0.45	–15.8***	0.41
Prep	1.39***	0.30		
n (unweighted)	64,794		64,794	
Censored	23,866		23,866	
	36.3%		36.3%	
Log Likelihood	–269,865		–269,875	

*** $p < .001$
** $p < .01$
* $p < .05$

Variable Definitions for Regressions

Agi	Adjusted gross income ($0000)
Est pay	Estimated tax payments ($000)
Baldue	Balance due (refund) upon filing ($000)
Mfj	Marital status indicator (0/1–Other/Married Filing Jointly)
Age	Aged indicator, primary or secondary taxpayer (0/1–Absent/Present)
Scha	Schedule A (Itemized Deductions) indicator (0/1–Absent/Present)
Schb	Schedule B (Interest & Dividends) indicator (0/1–Absent/Present)
Schc	Schedule C (Business Profit/Loss) indicator (0/1–Absent/Present)
Schd	Schedule D (Capital Gains/Losses) indicator (0/1–Absent/Present)
Sche	Schedule E (Supplemental Income/Loss) indicator (0/1–Absent/Present)
Schf	Schedule F (Farm Income/Loss) indicator (0/1–Absent/Present)
Schse	Schedule SE (Self – Employment Tax) indicator (0/1–Absent/Present)
IRA	Individual Retirement Account indicator (0/1–Absent/Present)
Keo	Keogh (H.R. 10) Retirement Plan Account indicator (0/1–Absent/Present)
F1040A	Form 1040A indicator (0/1–Absent/Present)
F1040EZ	Form 1040EZ indicator (0/1–Absent/Present)
Prep	Preparation mode indicator (0/1–Self-Prepared/Paid-Prepared)

TABLE IV
Regression Estimates
(Weighted Tobit)

Calendar Year 1988 Returns Filing during January–April 1989
Dependent Variable—Filing Date (Censored at 4/17/89)

	Self-Prep	Standard Error	Paid-Prep	Standard Error
Intercept	85.5***	0.66	84.1***	0.30
Agi	–0.241	0.15	0.160***	0.044
Estpay	0.488***	0.12	0.0757**	0.029
Baldue	3.57***	0.14	0.0797***	0.021
Mfj	–2.66***	0.59	–4.07***	0.32
Age	–5.34***	0.91	–1.91***	0.42
Scha	3.88***	0.80	3.44***	0.35
Schb	4.21**	1.67	6.48***	0.61
Schc	14.3***	1.47	11.3***	0.55
Schd	7.69***	1.35	6.52***	0.51
Sche	11.2***	1.13	10.5***	0.41
Schf	–5.19	3.13	–7.53***	0.82
Schse	6.52***	1.76	9.57***	0.60
IRA	1.43	1.07	–0.179	0.58
Keo	–5.51	4.74	1.56	1.73
F1040A	–11.9***	0.69	–13.1***	0.54
F1040EZ	–16.0***	0.72	–9.23***	1.08
n (unweighted)	24,861		40,933	
Censored	5,969		17,897	
	24.0%		43.7%	
Log Likelihood	–105,027		–160,781	

*** $p < .001$
** $p < .01$
* $p < .05$
Variable definitions: See Table III.

Table III. The coefficient estimates are virtually identical to those of the model with the preparer variable included, suggesting that the potential confounding effects of preparer choice are not large.

Table IV reports the results of estimating the basic equation separately for taxpayers who prepared their own return and for taxpayers who used a professional preparer. One striking difference emerges. Self-preparers are about 45 times more sensitive to the refund or balance due amount compared to taxpayers using professional preparers. A difference of only $280 (1000 divided by 3.57) is enough to induce a self-preparer to file a day earlier; much more than that is required for professionally prepared returns. Note in some cases the taxpayer may not have complete control over when the return is filed when a tax professional is involved.

Partitioning the sample by tax balance (i.e., refund or tax due) controls for one of the most important determinants of filing date. Table V shows that, when this is done, gross income is positively associated with later filing in both groups. A larger refund accelerates filing, but a larger tax due has only a very small effect on postponing balance due returns. The filing acceleration effect of being elderly is seen to be much larger for tax due returns; the reverse is true for the acceleration effect of filing a 1040A or 1040EZ return. We also see that earlier filing by farmers appears to be only for those that have a balance due. Inex-

TABLE V
Regression Estimates
(Weighted Tobit)

Calendar Year 1988 Returns Filing during January–April 1989
Dependent Variable—Filing Date (Censored at 4/17/89)

	Refund	Standard Error	BalDue	Standard Error
Intercept	78.4***	0.38	105***	0.41
Agi	0.150*	0.073	0.285***	0.057
Estpay	0.559***	0.052	–0.0165	0.037
Baldue	1.60***	0.12	–0.0750**	0.024
Mfj	–4.51***	0.38	0.684	0.42
Age	–1.92***	0.59	–15.8***	0.47
Scha	2.70**	0.46	7.87***	0.47
Schb	8.34***	0.95	0.920	0.70
Schc	13.1***	0.79	11.0***	0.74
Schd	4.30***	0.80	3.09***	0.59
Sche	13.2***	0.60	3.95***	0.50
Schf	3.82**	1.38	–21.0***	1.00
Schse	1.76	0.99	0.0737	0.74
IRA	1.96**	0.71	–2.54***	0.71
Keo	2.71	3.14	–4.50**	1.67
F1040A	–10.2***	0.46	–0.0967	0.71
F1040EZ	–12.5888	0.49	–0.814	0.83
n (unweighted)	38,151		27,643	
Censored	7,324		16,542	
	19.2%		59.9%	
Log Likelihood	–164,965		–83,983	

*** $p < .001$
** $p < .01$
* $p < .05$
Variable definitions: See Table III.

plicably, for returns claiming refunds, farming is associated with later filing. The relationship between filing date and tax-favored savings plans also appears to depend on tax balance—a positive association is detected only for IRAs and only for refund returns, while a negative association exists for both IRAs and Keoghs for returns with tax due. Most of these relationships continue to hold across tax balance after controlling for preparation mode, as shown in Table VI.

It is important to reiterate that, because both the use of a paid preparer and refund status are subject to taxpayer choice, these estimates are best viewed as descriptive rather than structural in nature. As an example, the presence of a complex tax status is likely to increase the likelihood of using a paid tax preparer, as well as filing date conditional on preparer status.

Longitudinal Analysis

To examine the persistence of procrastination over time, we analyzed returns from the 1979–1988 Statistics of Income Panel of Individual Returns, which is a part of the Ernst & Young/University of Michigan Tax Research Database. The Panel Files are subsets of the Individual Model Files that represent a simple random sample of individual income tax returns filed each calendar year. More importantly, although identifiers have been deleted and extensive safeguards have been taken to protect taxpayer confidentiality, each

TABLE VI
Regression Estimates
(Weighted Tobit)

Calendar Year 1988 Returns Filing during January–April 1989
Dependent Variable—Filing Date (Censored at 4/17/89)

	Paid-Prep BalDue	Standard Error	Paid-Prep Refund	Standard Error	Self-Prep BalDue	Standard Error	Self-Prep Refund	Standard Error
Intercept	105***	0.39	78.4	0.40	106***	1.11	79.7***	0.74
Agi	0.193***	0.047	0.255***	0.069	0.654**	0.23	0.0708	0.19
Estpay	0.0264	0.031	0.256***	0.047	–0.241	0.15	1.43***	0.16
Baldue	–0.0463	0.020	0.695***	0.11	–0.0327	0.16	5.83***	0.32
Mfj	–0.0926	0.40	–4.74***	0.44	1.53	1.12	–3.88***	0.64
Age	–14.4***	0.44	–0.803	0.62	–18.0***	1.30	–4.63***	1.10
Scha	7.77***	0.44	2.72***	0.48	7.34***	1.33	4.19***	0.90
Schb	1.78**	0.62	8.89***	0.92	–1.90	2.21	5.52**	2.08
Schc	10.5***	0.65	11.4***	0.77	11.4***	2.37	16.7***	1.69
Schd	2.10***	0.50	3.74***	0.79	5.97***	1.79	5.00**	1.68
Sche	4.11***	0.45	11.7***	0.59	3.05	1.66	15.3***	1.33
Schf	–20.0***	0.84	2.96*	1.24	–19.4***	3.95	5.18	4.04
Schse	–0.795	0.64	3.03***	0.94	3.47	2.44	–2.93	2.30
IRA	–2.65***	0.67	1.33	0.81	–2.08	1.94	2.94**	1.17
Keo	–4.75***	1.43	6.92*	3.22	–1.66	5.95	–2.76	6.09
F1040A	–5.78**	1.21	–9.25***	0.66	1.44	1.41	–8.74***	0.76
F1040EZ	–9.36**	2.41	–5.29***	1.29	0.428	1.57	–12.7***	0.79
n (unweighted)	21,376		19,557		6,267		18,594	
Censored	13,251		4,646		3,291		2,678	
	62.0%		23.8%		52.5%		14.4%	
Log Likelihood	–65,072		–82,122		–18,827		–81,765	

*** $p < .001$
** $p < .01$
* $p < .05$
Variable definitions: See Table III.

record contains a code based on the Taxpayer Identification Number that allows tracking the same taxpayers over time. The date is not available in the Panel, so we use the week the return was posted to the IRS Individual Master File (the posting "cycle") as a proxy for filing date. In the panel, our proxy for filing date is posting cycle, which ranges from 4 to 52. Both the median and mode posting cycle is 19 (the week of May 7th) in 1989, somewhat later than the filing date. That reflects the processing time at the Service Center, which varies by tax balance (refund/balance due). Although a less precise estimate of filing time, it is highly correlated with the calendar date that was used in the cross-sectional analysis. In the 1988 Model file, the Pearson correlation between the two is 0.911 ($p < .0001$).

Using current year returns from the Panel, we calculated the Pearson correlation of the posting cycle week from year to year for each return ID appearing in consecutive years. The posting cycle field is not available for 1979, so our analysis is for tax years 1980–1988 (returns filed during 1981–1989). The variation in sample size reflects varying sampling rates in the panel. Time of filing from year to year is positively correlated at 0.539 and ($p < .0001$ for all years), which indicates persistent behavior. Thus we conclude that many of the factors which we argue affect expected filing time, such as a household's discount rate and value of time, are, as one might expect, persistent.

Future extensions of our exploration of the temporal persistence of taxpayer filing behavior may help to explain some of the puzzling aspects of behavior we uncovered in the cross-sectional analysis.

VI. CONCLUSION

This paper presents the first empirical analysis of individual taxpayers' filing time. We find some evidence that is consistent with our model of the optimal filing date. Ceteris paribus, higher refunds are associated with earlier filing, complex returns are associated with later filing, and higher incomes (as a proxy for higher costs of time) are associated with later filing.

Our model cannot adequately capture the less comprehensible fact that millions of filers remit their taxes due before the filing deadline. These taxpayers, as a group, passed up $46 million in interest income in 1989. We do find, however, that this behavior is concentrated among the elderly, a group which on average has a lower value of time. It is possible that those who file early are averse to being in debt, fear forgetting or losing their return materials, or perhaps get utility from fulfilling their half of their contract with the government.[12]

Although our simple model leaves much of the variation in filing times unexplained, we believe that our results suggest that something akin to our theoretical model may be a good approximation to actual behavior. In our model people do not leave $100 bills lying around on the sidewalk forever. However, they may leave them there for some time while they wait for a moment when bending over to get the bill is relatively painless. Thus, we suggest that human behavior sometimes allows short-run profit opportunities to pass. They pass not due to individual irrationality or near-rationality, but rather because of the stochastic nature of individuals' opportunity costs of acting. As we document in our case of tax filing, in aggregate the amount of these foregone profits can be large: nearly a billion dollars in interest was foregone by the group of taxpayers who were due refunds and chose to file at the deadline in 1989.

A more sophisticated analysis would allow for the tax filing time to be jointly determined with choice of preparer status and refund status. The latter connection is particularly intriguing because it involves (as does the filing time decision) foregoing interest by remitting taxes earlier than necessary. We leave to future research the task of an integrated analysis of tax withholding (and estimated tax payments) and filing time.

12. The elderly, for example, receive far more in benefits from the Federal government than they currently pay for (and lifetime resources received also far outweigh lifetime payments).

REFERENCES

Akerlof, George A. "Procrastination and Obedience." *American Economic Review*, May 1991, 1–19.

Cordes, Joseph J., Harvey Galper, and Sheila Nataraj Kirby. "Causes of Overwithholding: Forced Savings or Transaction Costs? An Indirect Test of the Behavioral Life-Cycle Model." Manuscript, George Washington University, October 1988a.

_____. "Causes and Implications of Overpayment of U.S. Individual Income Tax Liabilities: An Empirical Analysis." Manuscript, George Washington University, October 1988b.

Engberg, John B. "The Impact of Unemployment Benefits on Job Search: Structural Unobserved Heterogeneity and Spurious Spikes." Manuscript, Carnegie Mellon University, School of Urban and Public Affairs, 1991.

Laibson, David. "Self Control and Saving." Manuscript, Massachusetts Institute of Technology, 1994.

Lippman, Steven, and John J. McCall. "The Economics of Job Search: A Survey, Part I." *Economic Inquiry*, June 1976, 155–89.

McCall, John J. "Economics of Information and Job Search." *Quarterly Journal of Economics*, February 1970, 113–26.

Mortensen, Dale T. "Unemployment Insurance and Job Search Decisions." *Industrial and Labor Relations Review*, July 1977, 505–17.

Wolpin, Kenneth I. "Estimating a Structural Search Model: The Transition from School to Work." *Econometrica*, July 1987, 801–17.