Consumption smoothing among working-class American families before social insurance

John A. James
Department of Economics
University of Virginia

Michael G. Palumbo
Board of Governors of the Federal Reserve System

Mark Thomas
Department of History
University of Virginia

Abstract: This paper examines the saving decisions of a large sample of turn-of-the-century working-class American families. We decompose each family's reported income into permanent and transitory components and then estimate marginal propensities to save from each component. Marginal propensities to save out of transitory income are large relative to the propensities based on permanent income, though the former lie much below one and the latter much above zero, remarkably similar to results based on contemporary data sets Smoothing appears to have been primarily at medium rather than low frequencies, more consistent with precautionary than with life-cycle motives.

JEL Classification Codes: D91, E21, N31.
Keywords: Unemployment, uncertainty, permanent income hypothesis, precautionary saving.

1. Introduction

Around the turn of the twentieth century, many Americans lived and worked in an environment of considerable economic uncertainty. Industrial accidents presented significant risks to many (see Kantor and Fishback, 1996), as did illnesses, but even more pervasive was the risk of unemployment, which was much greater during the period before World War I than it has been since World War II. Not only was the natural rate higher, but so too was the cyclical sensitivity of unemployment (James and Thomas, 2003, 2004). Moreover, the incidence of unemployment was more widespread, implying that a greater proportion of workers had need for precautionary action than today. Unemployment in this historical period had its predictable elements—the availability of work followed strong seasonal influences, for one thing—but loss of work also resulted from much less predictable factors. Business cycle downturns during the late nineteenth and early twentieth centuries were on average more serious than they had been before or have been since. The severity of business cycles had increased dramatically from the period before the Civil War to the one after (James, 1993)—indicating an increasing need for precautionary behavior—and this in a period before the rise of governmental institutions designed to take the sting out of being put of work. Workers in the late nineteenth century were essentially dependent on their own devices to combat income uncertainty. It is perhaps ironic that the expansion of social insurance after World War II coincided with a moderation of unemployment volatility.[1]

Alexander Keyssar, in his well-known study of unemployment in Massachusetts, observes that employment for workers in this period was 'chronically unsteady' (1986, p. 59). Even within the business cycle he stresses the great diversity of individual experience: 'The incidence of joblessness during depressions was always checkered, erratic, variegated.' Moreover, 'a majority of the working class found that the threat of unemployment remained palpable even when business was good' (1986, pp. 55, 58). Substantial negative shocks to

---

[1] Long run changes in the degree of volatility of national output and unemployment in the U.S., of course, have been a controversial issue (see Romer, 1986; Weir, 1992). However, James and Thomas (2004) document a decrease in the cyclical response of unemployment from the pre-World War I period to that post-World War II.

household income therefore would have been common and, especially for lower income families, presented potential economic disasters. Workers faced such risks without assistance from any sort of public safety net—no unemployment compensation, no worker's compensation for accidents, no sickness coverage. Few workers had access to formal credit markets during this time to enable smoothing adverse income shocks by taking temporary loans against future income or accumulated assets; there were no credit cards, no installment plans, no home equity lines of credit (see de Long and Summers, 1986). Commercial banks, influenced at least in theory by the real bills doctrine, limited themselves to commercial rather than personal loans, and rarely included working men among their clientele.

How then did working families meet the challenges of income shocks before 1914? Borrowing from friends and relatives, running up store credit, pawning jewelry, clothing and other personal property at the local hockshop, letting rents fall into arrears, and, most notably, dipping into accumulated savings at home (under the mattress, in the jar on the mantelshelf) or in a savings bank may well have the most important means to maintain consumption in the wake of sickness or unemployment, especially for the most vulnerable.[2] The lower paid, unable to build up a nest egg of cash or having already exhausted it, probably were the most likely customers for the pawnbroker—whether on a regular or an occasional basis [3]—and may also have been recipients of such limited private charity as was available in this period to support 'the less well-off.' Those with more skills and higher (and perhaps more regular) incomes may have been able to take advantage of more formal institutions, via their membership of mutual benefit societies and their maintenance of mutual savings bank accounts[4]. Note however that even for these

[2] Out of 357 families who declared to BLS survey takers in 1902 how they met an income shortfall (expenditure in excess of income) in the previous year, 45% obtained credit, 26% depleted savings (whether held in banks or at home), 4% borrowed money, and the remainder mortgaged or sold property or real estate. Unfortunately, the survey did not record the amounts involved (US Commissioner of Labor, 1904).

[3] Note that the average pawnbroking loan was small--$4.14 according to a federal survey in 1897 (Patterson, 1899, p. 274).

[4] About a quarter of the working men in our sample belonged to some kind of benefit society (some belonged to more than one) offering sickness, accident, life or burial (but not old-age) insurance. A BLS survey of 2,567 families in 1902 indicated that one-third held savings bank accounts; about half of those families who had a surplus

households, it would have been necessary to utilize informal agencies to smooth consumption, given that most benefit societies required a minimum period before benefits could be paid, and that savings banks required notice of withdrawals and did not permit borrowing.[5] All working households faced credit constraints, but some more than others.

A similar environment of unpredictable incomes and weak formal credit markets exists for low-income households in developing countries today. Nonetheless, a number of recent studies have found that, even without access to formal insurance and credit markets, such households have generally been able to smooth their expenditures in the face of large and frequent income shocks (e.g., Deaton, 1990; Paxson, 1992; Townsend, 1995). However, almost all of the households studied in developing countries rely heavily on agricultural production for their earnings and can employ mechanisms that were not available to American industrial workers a hundred years ago, such as crop diversification and the maintenance of buffer stocks of commodities.

Does it follow, therefore, that American working-class families responded to uncertain earnings prospects through different private saving patterns than has been observed among peasant farm households today? Or does it mean that they were unable to smooth consumption in the face of volatile and unpredictable incomes? The first part of the paper is an econometric analysis of the role played by private saving to buffer volatile income experiences a century ago. The econometric results suggest saving patterns consistent with intertemporal optimization, but different specifications of models incorporating intertemporal optimization imply different horizons over which consumption is smoothed—from over the lifetime in the 'pure' life-cycle model to relatively short periods in buffer-stock models (e.g., Carroll, 1997). In the second part of the paper we investigate the smoothing horizon and argue that the observed saving patterns

---

of income over expenditure in the previous year deposited their savings in the bank, while 36% kept the excess at hand, 8% invested in real estate or put money in a building and loan, and 5% paid off existing debts. The survey did not record the amounts involved (US Commissioner of Labor, 1904).

[5] See the evidence of George V. Cresson, President of the Manufacturers' Club of Philadelphia, to the Industrial Commission, December 18, 1900: 'You cannot draw the money out [of a savings bank] except after so many weeks' notice' (Industrial Commission, XIV, 271).

are consistent with precautionary rather than life-cycle motives.

The plan of the paper is as follows. The second section describes the most comprehensive micro-level database to date on saving, income and employment covering American workers in the Gilded Age. By merging information from over twenty cross-section surveys based on nearly identically-worded questionnaires, our database includes information from more than 32,000 working-class families interviewed between 1879 and 1909. In the third section, we describe an econometric methodology, similar to that proposed by Paxson (1992), for decomposing annual income realizations into permanent and transitory components for each worker in the sample. The fourth section contains the primary empirical results of the paper based on estimated marginal propensities to save out of permanent and transitory income. Our empirical analysis presents strong empirical evidence that working-class American households used their own saving to smooth consumption in the face of volatile incomes during the late nineteenth and early twentieth centuries, much as contemporary American and European families, as well as farmers in developing countries appear to do. The fifth section examines the horizon at which households tried to smooth consumption. As almost all recent studies find as well, the historical evidence rejects strict specifications of the life-cycle or certainty equivalence hypotheses but seem quite in line with what one might expect in specifications which allow for precautionary saving behavior. Section 6 summarizes and concludes.

2. Saving data from a series of worker surveys

State-level bureaus of labor statistics published more than one hundred surveys of wage-earning workers during the late nineteenth and early twentieth centuries. The surveys focused on economic conditions facing American workers, employed mostly in the non-farm sector, and the living conditions of their families. Some surveys concentrated on workers in specific industries (vehicle manufacture; iron production; furniture), holding interviews at the workplace (the Michigan model); others polled representative samples of workers across industries by mail (the Kansas model). Although the information collected varied from survey to survey, most used similarly worded or identical questions, following the example of Carroll D. Wright in Massachusetts, who began surveying workers during the early 1870s. Individual survey responses were invariably published in full, without editorial embellishment or alteration once accuracy and consistency were checked. Each survey covered a different cross-section of workers; none followed the experiences of individual workers or families across more than one year. It is not possible therefore to construct a true panel from these sources; our database consists of merged independent cross-section surveys.[6] As Table A-1 indicates, the database includes 32,168 workers from 6 states interviewed during 22 different annual periods from 1884 through 1909. Note, however, that nearly all (91 %) of the families in our sample resided in Kansas or Michigan; most of the observations used here (71 %) were drawn from the 1890's.

We limit our analysis in this paper to those surveys which report information on income, saving and/or expenditures and days of loss of work during the previous year. Additionally, we know the skill level for each worker, as well as his age, industry, state of residence and the year of the survey.[7] Income in this paper refers to annual family income, deflated to 1900 prices, combining income from all sources. A few surveys report labor earnings disaggregated by earner (or, at least, separately for primary wage-earners and all others in the family), but most do not.[8] The database employs two different measures of annual saving.[9] Some households reported last year's saving directly (answering a question like 'How much of your income did

---

[6] The data set is discussed in more detail in James, Palumbo, and Thomas (2005).

[7] Our working database excludes the few women surveyed, as well as the few men younger than 15 years of age or older than 75.

[8] None of the surveys separate interest income from labor income, but the former category is unlikely to be a major contributor among workers included in this sample.

[9] Note that the surveys clearly asked respondents about the annual flow of saving out of income during the previous year, rather than about their accumulated stocks of assets or net worth. Some of the respondents, naturally, owned their homes. During this period, home mortgages typically involved short-term contracts in which only interest payments occurred during the loan's duration with a balloon payment of principal due at maturity. Thus, paying down the loan principal at maturity must have accomplished through prior saving and, we have argued (James, Palumbo and Thomas, 2005), our saving measures seem likely to include changes in home equity. Also, the range of consumer durables available to working-class families at this time was quite limited. Consequently, potential distortions in measuring saving due to neglecting consumer durable expenditure should be quite small.

you save last year?’); others recorded total family expenditures, from which saving may be calculated as the residual from annual income.[10] We designate the first variable as ‘reported saving,’ and the second as ‘calculated saving.’

Figs 1 and 2 display empirical density and distribution functions for reported and calculated saving based on the pooled cross-section survey data. The figures show clear differences between the two distributions. Surveys that measured saving directly only recorded additional money set aside (positive values for reported saving); otherwise, entries were left blank. After analysis of the data, we concluded that 'no response' generally indicated zero (or negative) saving.[11] Thus, Fig. 1 shows some possible effects of left-censoring at zero—a large spike in the distribution function at the censoring point—as well as other smaller spikes at ‘round’ numbers.[12] Calculated saving, on the other hand, takes both positive and negative values. Fig. 2 therefore shows a less skewed distribution of saving, and a smaller spike occurring exactly at zero dollars.

Left-censoring in reported savings ought to produce a higher average level of savings than that based on calculated saving, which takes some negative values. As it happens, the mean values of the variables are reversed relative to expectations—average reported savings equals \$57.74 per year, compared to \$98.85 for calculated saving. Some of this gap can be traced to differing compositions of the two samples of families, in terms of their ages and income levels. The subsample for whom saving is reported directly in the surveys is both younger (about 32 versus 37 years, on average) and poorer (annual real income averages \$543 versus \$579) than the subsample for which we calculate saving as the discrepancy between reported income and

---

[10] Direct savings were recorded by 26,070 families, while savings were calculated as a residual for 12,012 families (see Tables A-1 and A-2).

[11] One survey, covering vehicle workers in Michigan during 1896, recorded three responses to the question ‘amount saved last year?’ – a positive value, ‘none,’ and a blank response. Of the 2,787 survey respondents who do not report positive saving (out of 3,776 total observations), 2,576 explicitly answer ‘none’; only 211 have a blank response.

[12] For consistency, we artificially censored the few observations for which we observe negative values for reported saving. These come from Kansas surveys in which respondents answered ‘Did you save or run a deficit last year?’

expenditures. Correcting for compositional differences reduces the gap between the two groups considerably—we calculate that average calculated saving would only have equaled $79.72 if its sample displayed the age and income composition of the 'reported saving sample'—but the sign on the gap is still counter-intuitive.

A further explanation lies in the tendency for families either to understate their reported saving or to understate their reported expenditures relative to their reported incomes. This can be deduced by examining reported and calculated saving among the 7,957 families for which both variables are available. Calculated and reported saving are highly correlated in this subsample. A regression of calculated saving on reported saving yields a slope coefficient equal to 0.9930 ($R^2 = 0.53$), with an intercept of $27.50. This result suggests that households may have taken a question on saving to mean literally, 'money put away,' rather than accidental differences between income and expenditure; or that working households were not always able to account for (or chose not to report) all their expenditures; or some combination of the two.

This divergence however turns out to have virtually no consequence for our empirical analysis. Given the strong linear relation between reported and calculated saving, it should not be surprising that none of the econometric results in this paper are sensitive to the choice of dependent variable. As we report below, Tobit equations estimated using reported saving yield nearly identical slope coefficients as OLS regressions based on calculated saving, though the estimated intercepts differ between the specifications.

Table A-2 presents mean income, mean consumption and the variance of expenditure relative to the variance of income for workers in our data set who report information on annual saving directly (grouped by age, by skill and by industry) in the first panel. Similar information is shown in the second panel for those workers who directly report income and expenditures (i.e., for those whose savings have been calculated by us). Both clearly indicate the tendency for variability in family income to exceed variability in expenditure, regardless of classification. In the subsample for which calculated saving is available (panel II) the cell variance of consumption is generally at least 30 % less than that of income; where saving was reported directly (panel I), the difference is somewhat smaller—expenditures vary about 20 % less than incomes, on average.

An interesting pattern follows from comparing mean income and consumption between savers and nonsavers within each cell. In virtually every cell, savers earn higher incomes than nonsavers, but average consumption levels between the two groups are very similar. The few cells in which consumption differences are relatively large all suffer from relatively small cell sizes. Among families in the calculated saving subsample, mean consumption levels differ only by 8 % between savers and nonsavers, but income levels are about 34 % greater among savers on average. This striking result suggests that saving might have responded largely to income 'surprises' among these families, thereby motivating the empirical strategies to be described next.

3. Estimating permanent and transitory components of annual income

A key implication from theories of household saving based on intertemporal optimization is that marginal propensities to save (mps) or to consume (mpc) differ by composition of income. The empirical strategy for examining saving behavior employed here involves making explicit comparisons of the marginal propensities to save out of permanent and transitory income to discern motives for saving. 'Keynesian-saving' families, whose spending simply is a function of current income, ought to have marginal propensities to save that apply equally to all components of annual income. On the other hand, if families are guided by a 'certainty equivalence' decision rule (or according to the 'strict permanent income hypothesis' in Deaton's (1992) terminology), permanent differences in income levels will not affect observed saving levels, while transitory income shocks will affect saving decisions dollar-for-dollar.[13] Finally, recent theoretical models based on intertemporal optimization with unpredictable family incomes, 'impatient' and 'prudent' consumers (both characteristics of household utility; see Carroll, 1997 or Deaton,

[13] A formal derivation for optimal consumption and saving rules under certainty equivalence, or the strict permanent income hypothesis, can be found in Pistaferri (2001); Paxson (1992) includes an informal discussion of similar results.

1992), and either with (Deaton, 1991) or without (Carroll, 2001b) liquidity constraints suggest small, but nonzero, marginal propensities to save out of permanent income, and larger marginal propensities (less than one) out of transitory income.

This approach to econometric analysis of household saving behavior thus requires the decomposition of realized family income each year into its permanent and transitory parts. After Paxson (1992), rather than estimating transitory income simply as a residual from a regression equation for annual income on permanent family and worker characteristics, we use survey information to measure it more directly. In her study of saving behavior among Thai rice farmers, Paxson uses deviations in rainfall from historical averages to measure transitory income shocks for small geographic regions in Thailand. Our application focuses on shocks to time spent out of work among primary earners in working-class families. Unexpected days lost from work (whether from low-frequency job loss, high-frequency inabilities to find work, accidents, or sicknesses) would not have directly affected family expenditures, but would have translated into important income shocks to families in our database, which theory suggests should flow into changes in savings levels. A deviation in reported workdays lost from its predicted value, therefore, provides a measure of transitory income shock realized by each family in our database.[14]

Decomposing annual income into its permanent and transitory components requires us first to predict annual workdays lost for each survey respondent. The Appendix contains an algebraic representation of the model which we describe verbally here. We first estimate a regression for annual workdays lost as a function of each respondent's age, state of residence, and skill category interacted with industry of employment. We next use the regression-fitted value to estimate the predictable number of workdays lost for each observation, based on all

[14] Instability in incomes may arise from volatility in either wage rates per day or in the number of days worked. Analysis of variance of income instability among 516 Pennsylvania workers who reported their income and employment for 1890, 1891, 1892, and 1893 indicates that days lost accounted for 64 % of total variance; wage variability accounted for 31%; the covariance of employment and wages accounted for the rest. This is consistent with the findings of, inter alia, Sundstrom (1990) and Hanes (1993) which find widespread evidence of wage rigidity during the 1890s. Given the high correlation of wages and prices in this period, the contribution of real wage variability to total income instability would have been even smaller.

available years in the sample. The regression residual then estimates the unexpected shock to employment experienced by each worker during the previous year. Our specification is quite flexible and, because all the explanatory variables are categorical indicators, the regression effectively defines the predictable number of workdays lost to be the average among all workers of a particular type. We assign workers to types, or cells, defined by four age categories, four skill groups, six industry groups, and six state indicators. Rather than report all the estimated regression coefficients, which are cumbersome to interpret, column (1) of Table 1 shows average workdays lost during the previous year by categories of worker type (state groups not being reported here for conciseness). The regression results show clear differences in average workdays lost among different groups of workers.

The unexpected component of annual workdays lost is defined as each individual survey respondent's deviation in reported workdays lost during the previous year from the average workdays lost among members of his type. The second and third columns of Table 1 show estimated dispersion in unexpected workdays lost during the previous year obtained from this procedure, as measured by the interquartile range by worker types. Note that, by construction, unexpected workdays lost must average zero for each worker type. Fig. 3 shows the density and distribution functions for unexpected workdays lost realized by workers in our sample estimated according to these procedures.

Having in hand an estimate of unexpected workdays lost, we now must translate that variable into a measure of transitory income during the previous year for each family in the sample. This involves estimating a regression for annual family income on unexpected workdays lost by its primary wage-earner during the previous year and the same set of explanatory variables used in the workdays lost equation to measure predictable income movements. In addition, unexpected workdays lost is interacted with the set of categorical variables to allow the price of a day lost to differ across families. The income regression, estimated using the entire sample of 32,168 families, yields an adjusted R^2 of 0.40.

Permanent family income for each sample household is calculated using the fitted value from the age-, survey- and skill-by-industry indicator variables (and their estimated coefficients) in the family income regression. Column (4) of Table 1 reports average family income, which equals average permanent income by definition, across worker types defined in the regression. In turn, transitory income for each family is calculated as the product of estimated unexpected workdays lost during the previous year and its interactions terms with their respective estimated coefficients. The distribution of transitory income by category of worker is summarized in columns (5) and (6) of Table 1. Finally, the difference between income actually reported by each family in the survey and estimated permanent and transitory income levels can be computed. Following Paxson (1992), we call this residual 'unexplained annual income.'

To summarize our empirical methods to this point, respondents in our sample are assumed to have used average incomes earned by other families 'like themselves'—a group defined by their age, state of residence, as well as their occupational skill and industry classification—to predict their annual incomes. The estimated coefficient on unexpected workdays lost in the family income equation allows us to derive a dollar-measure of transitory income. Transitory income shocks, therefore, average zero across all workers of a given type—or within each cell—but vary among the individuals within each cell. We then use our estimates of permanent and transitory income as regressors in an equation for family saving decisions, as is described next.

4. Explaining household saving with permanent and transitory income

We take household saving to have been a linear function of the components of realized annual income:[15]

[15] Experimentation with various nonlinear functions (income categories; splines; quadratic/cubic terms) did not indicate substantial or significant departures from linearity in these data. Paxson (1992) uses time-series data on the standard deviation of regional rainfall to estimate the variance of annual income among her sample of Thai farmers, which she also includes in her saving equation (1). Not having pure time-series variation, we use an alternative, intragroup variation, in estimating our equation (2) (we provide more justification for this in Section 5), but in our case here there are 220 cells, so to simplify matters we do not reproduce Paxson's equation (1) specification exactly here. This simplification has no effect on the tenor of the results for the coefficients for permanent and transitory

$$S_{i(t)} = \alpha_0 + \alpha_1 \hat{Y}^P_{i(t)} + \alpha_2 \hat{Y}^T_{i(t)} + \alpha_3 \hat{Y}^U_{i(t)} + \epsilon^S_{i(t)} \quad (1)$$

where the subscript denotes an observation on family i surveyed during year t and the explanatory variables, respectively, are permanent income, transitory income and unexplained income components, as estimated according to the methods described in section 3. Models that assume that incomes are predictable over the lifetime (the 'perfect foresight' model) or that assert that consumers behave as if their incomes were certain ('the certainty equivalence' model) predict that the marginal propensity to save out of permanent income, α_1, is zero, while the marginal propensity to save out of transitory income, α_2, approaches one. Precautionary savings models that assume that households face uncertain incomes and respond 'prudently' by building up an 'emergency reserve' are consistent with a positive, but small marginal propensity to save out of permanent income and a marginal propensity to save out of transitory income that is large, but less than one ($1>\alpha_2>\alpha_1>0$).[16] Simple 'Keynesian' consumption rules would predict no difference between consumption out of permanent and transitory income ($\alpha_1=\alpha_2$), while 'hand-to-mouth' behavior, in which families spend their current incomes each year, is consistent with $\alpha_1=0$ and $\alpha_2=0$.

4.1 Saving regressions based on respondent-level data

Tables 2 and 3 present the primary estimation results for several model specifications. Panel I of Table 2 shows results based on Tobit equations using the left-censored (at zero) reported saving as the dependent variable, while panel I of Table 3 contains OLS estimates when calculated saving is the dependent variable. Both tables report consistent standard error estimates based on

income reported in the text.

[16] See, e.g.., Carroll (1997, 2001a, 2001b). Browning and Lusardi (1996, p. 1808) note that precautionary models are generally 'compatible with a much richer variety of short-run and life-time consumption patterns than is suggested by the CEQ [certainty equivalence] model.' Deaton (1991) and Carroll (1997, 2001a, 2001b) add impatience, in the sense that households would prefer to consume today, rather than tomorrow. In that case, the size of the 'emergency reserve' fund will be limited by the 'tug-of-war between impatience and prudence' (Carroll, 2001a, p. 24). Note that Deaton (1991) makes no distinction between permanent and transitory income.

a complete resampling bootstrap procedure using 200 replications of the three-stage estimation procedure.[17]

The first column in each table presents the estimated results from the baseline specification; columns (2) through (4) report results from a few of the many alternative model specifications with which we experimented. In column (2) we constructed a new set of survey weights to generate hypothetical coverage of 10,000 respondents in each survey and then reestimated the saving equations by Tobit and OLS respectively, so as to reduce the influence of the Michigan data relative to those from other states. The results indicate that our basic results are not driven by unrepresentative behavior of upper mid-western households. The results in column (3) employ another set of sample weights, by estimating the saving equations using only respondents from Kansas and Michigan, which together provide more than 90 % of our original sample. Once again, this revision does not alter the basic parameter estimates relative to the baseline case. In column (4) we include a quadratic term in age of the household head in the regression equation, to allow for the possibility that longer-term life-cycle factors may also have had an influence on saving other than through permanent and transitory income. The inclusion of the age and age-squared variables has little effect on the estimates of the other variables. Both equations show saving increases with age, other things equal, although the shape of the age functions differ between the two data sets.[18]

The parameter estimates reported in Tables 2 and 3 are remarkably similar, regardless of whether saving is measured by reported or calculated saving, or how the data set is weighted. The coefficients show that both permanent and transitory components of annual income influence household annual saving, but with significantly different marginal effects. Marginal

---

[17] We repeatedly estimated all three equations of our empirical model— workdays lost, family income, and saving as a function of permanent and transitory income components—using 200 bootstrap random samples, then constructed standard errors for the saving equation regression coefficients by computing standard deviations across the replicated samples. OLS standard errors are incorrect because of the presence of constructed variables (the three income components) in the saving regression equations.

[18] We also have investigated some further additional specifications of the saving equation, the results of which have been omitted from the tables to conserve space. First, by interacting dummy variables for three age categories with all three income components, we estimated saving equations with age-specific values of α_1, α_2 and α_3. The results indicate quite different reported-saving equations for respondents aged less than 25 years compared to older respondents. Among the youngest group of respondents, marginal propensities to save out of permanent income were nearly as large as the marginal propensities to save out of transitory income. These results were not replicated in regressions based on calculated saving, which raises questions about the robustness of age-specific savings behavior. Second, we explored saving equations in which skill indicators or industry indicators were interacted with income components. These regression specifications, however, did not yield substantially different behaviors among different groups of respondents.

propensities to save are much larger for transitory income than for permanent income: the point estimates range between 0.26 and 0.35 for saving out of permanent income, and between 0.47 and 0.59 for saving out of transitory income. The test statistics for the null hypothesis of equality ($\alpha_1=\alpha_2$) are reported at the bottom of Tables 2 and 3; in every specification the null is rejected at conventional levels of significance. The estimated propensity to save out of unexplained income, a mixture of permanent and transitory income, should reflect the influence of both components. In all save one case, estimates of α_3 lie between those of α_1 and α_2.

4.2 Unobserved heterogeneity and its potential impact on respondent-level regressions

The regressions described so far would produce biased estimates of α_1 and α_2 if regression errors for family saving are correlated with permanent and transitory income components at the respondent level. A potential problem arises because permanent and transitory components of income are variables constructed from regression fitted values, rather than being directly measured. This raises the vexing issue of unobserved heterogeneity. What if, for example, respondents knew more about their own work prospects than we attribute to them *and* if those who face the worst work prospects differ in their saving propensities systematically from those with the best work opportunities? In that case, our regressions would incorrectly attribute the effects of permanent and transitory income on family saving decisions.

To be concrete, consider the empirical implications for respondent-level regressions of the following type of unobserved heterogeneity with respect to labor supply and family saving behavior. Suppose, for simplicity, that our sample combines two types of workers: Type A respondents, who are hard-working and foresighted—they tend to work a lot of days during the year and to save relatively large portions of their incomes; and Type B respondents, who are myopic and less industrious. Further, suppose all groups of workers (defined as combinations of age, skill, industry, state and year categories in reference to our 'workdays lost' and income regressions) contain some respondents of both types. Type A respondents can be expected to report below (-group-) average workdays missed during the previous year, while Type B respondents report above-average workdays lost. Additionally, Type A respondents would save more than average, while those of Type B would pull the average down.

In this case, our regression procedure would overestimate permanent income among Type B workers and underestimate it among those of Type A, while overestimating transitory income among Type A workers (the estimates would be too large and positive, on average) and underestimating it among Type B workers (theirs would tend to be 'too negative'). Incorrect estimation of permanent and transitory incomes due to unobserved heterogeneity cause respondent-level regressions, such as those reported in panel I of Tables 2 and 3, to underestimate α_1 (the estimated regression line will be flatter than the true relationship), and to overestimate α_2 (the estimated regression line will be too steep).

In the absence of panel data, with multiple observations of lost workdays, income and saving for each respondent over a period of time, which would allow models with individual-specific effects to be estimated, we use instrumental variables estimation to wash out the potential distorting effects of unobserved heterogeneity. In the first stage family income is regressed on categorical variables for age, skill, industry, and state. The fitted values from this regression are taken as our estimate of permanent income, $\hat{Y}^P$, while the residual is taken as the estimate of transitory income, $\hat{Y}^T$.[19] The second stage is an instrumental variables micro-level regression. We use group-by-year interactions as instruments for $\hat{Y}^T$, the time variation of groups being crucial for identification. Since we cannot rely on theory to guide our group definitions,[20] we present results from two alternative specifications—in the first based on birth cohort, year, and cohort-by-year groups; the second, on age, skill, industry, state, year, age-by-year, skill-by-year, and industry-by-year groups.[21]

---

[19] Note that in this case we do not use unexpected days lost in directly constructing an estimate of transitory income, leaving unexplained income as the residual. The tenor of the reported results which follow however does not vary from the method used in calculating unexplained income.

[20] There are however some criteria for choosing good instruments. We would want membership in the group to be exogenous to the saving decision, conditional on permanent and transitory income, and for group membership to have explanatory power for income to get useful variation in permanent income across groups. Groups also should have different time paths for income surprises, that is, the interactive terms with year should have explanatory power for transitory income.

[21] Small cells (fewer than 44 observations, an apparent natural break point) are excluded to lessen the impact of outlying observations on the grouping results.

Estimated coefficients are reported in panel II of Table 2 for the Tobit regressions based on reported annual saving, and in panel II of Table 3 for the regressions based on calculated annual saving. If unobserved heterogeneity were an important source of bias in our respondent-level regressions, the group-level regressions ought to yield substantially larger marginal propensities to save out of permanent income and smaller marginal effects from transitory income. The instrumental variables regressions do not however reveal larger estimates of α_1 or smaller estimates of α_2. The estimated marginal propensities to save out of permanent income from the instrumental variable grouping regressions are certainly no higher on average than the baseline parameter estimates shown in panel I of Tables 2 and 3. Furthermore, the instrumental variables regressions yield substantially larger point estimates for the marginal propensity to save out of transitory income than the baseline parameter estimates in panels I, reinforcing our interpretation. This pattern holds consistently, for both sets of instruments and for both definitions of the dependent variable (reported and calculated saving).

These alternative specifications, therefore, do not support the concern that unobserved heterogeneity with respect to labor supply and saving behavior generates incorrect inferences from respondent-level regression analysis. We conclude that small marginal propensities to save out of permanent income and large propensities to save out of transitory (shocked) income characterize important behavioral patterns guiding saving decisions among working-class American families a hundred years ago, rather than differences in attitudes toward work and saving at the individual level.

5. Smoothing at different frequencies

The empirical results reported in sections 3 and 4 are inconsistent with simple models of savings behavior by working-class American households at the end of the nineteenth century. The finding that both α_1 and α_2 are statistically significantly greater than zero shows that families did not live 'hand-to-mouth,' while the statistically significant difference between α_1 and α_2 in all versions indicates that households did not follow simple Keynesian consumption rules.

5.1 Smoothing over the long-run

The econometric results reported in Tables 2 and 3 are also inconsistent with a strict life-cycle interpretation. The estimated values of α_2, the marginal propensity to save out of transitory income, are consistently statistically significantly smaller than one and those of $\alpha_{1,}$ the marginal propensity to save out of permanent income, greater than zero, their values in a certainty equivalence model (see the test statistics at the bottom of panel I).[22]

These econometric results may, however, hide an asymmetry in the response of households to positive and negative income shocks. For positive shocks, households would save in a manner consistent with the life-cycle model; but for negative shocks, liquidity-constrained households may have responded by lowering current consumption and saving nothing. The estimated less than one coefficients on transitory income may then have been the result of such an asymmetric response to positive and negative shocks due to liquidity constraints. To test this, we reestimate equation (1), distinguishing between positive and negative transitory income. The results from a Tobit regression with reported saving as the dependent variable and bootstrapped standard errors in parentheses are:

$$\underset{}{S_{i(t)}} = \underset{(6.062)}{-207.674} + \underset{(0.010)}{0.3009}\,\hat{Y}^{P}_{i(t)} + \underset{(0.046)}{0.6694}\,\hat{Y}^{T}POS_{i(t)} + \underset{(0.027)}{0.5515}\,\hat{Y}^{T}NEG_{i(t)} + \underset{(0.010)}{0.4192}\,\hat{Y}^{U}_{i(t)}$$

Those, in turn, from an OLS regression with calculated saving as the dependent variable are:

$$S_{i(t)} = \underset{(6.663)}{-76.860} + \underset{(0.015)}{0.3001}\,\hat{Y}^{P}_{i(t)} + \underset{(0.043)}{0.5854}\,\hat{Y}^{T}POS_{i(t)} + \underset{(0.022)}{0.4547}\,\hat{Y}^{T}NEG_{i(t)} + \underset{(0.012)}{0.4152}\,\hat{Y}^{U}_{i(t)}$$

[22] Several empirical analyses of contemporary family saving behavior (e.g., Flavin (1991), Paxson (1992), Alessie and Lusardi (1997) and Pistaferri (2001)) also reject the strict permanent income or certainty equivalence (CEQ) model (Deaton 1992; Browning and Lusardi 1996). Our estimated saving propensities out of permanent income for late-nineteenth/early-twentieth century American working-class families are quite similar to Paxson's estimates (1992), which ranged between .25 and .28 based on comparable saving definitions, for Thai rice farmers. On the other hand, Paxson's estimated marginal propensity to save out of transitory income, .74 to .75, are rather larger than our estimates. Furthermore, our estimated parameters for permanent income movements are very similar to estimates of the change in saving due to changes in permanent income (.18 to .25) from the excessive sensitivity of consumption studies done by Flavin (1991) for American families in the late 1960's, and by Alessie and Lusardi (1997) for Dutch families in the mid 1980's. Pistaferri's analysis of the saving decisions observed among contemporary Italian families yields marginal propensities to save out of permanent movements in income around 0.16 and out of transitory income shocks between about 0.50 and 1.20. Thus, despite differing specific empirical approaches and microdata from quite different economies, all five of these papers support the conclusion that families use saving to smooth consumption expenditures at a relatively high frequency (such as over the business cycle), but also that saving responds more to permanent income movements than it 'should', where reference is made specifically to behavior from the permanent income hypothesis.

The estimated coefficients on negative transitory income are less than those on positive ones in both specifications, but the test results are mixed. The null hypothesis of equality of the coefficients of positive and negative transitory income is rejected for the calculated saving variable (F(1,12007)= 9.64 with a significance level of 0.0019), but not for reported saving (F(1,26066)= 0.12 with a significance level of 0.7275). But the results provide no comfort for the certainty equivalence model. The positive transitory income coefficient still remains significantly less than one (F statistics of 70.57 and 79.35 respectively for reported and calculated saving, both with significance levels of 0.000); the estimated propensities to save out of permanent income remain substantial in magnitude and statistically significantly greater than zero; and the difference between the estimated coefficient on permanent income and the (smallest) estimated coefficient on transitory income is statistically significant.

These findings challenge the emerging historiography about savings behavior in late nineteenth-century America, which emphasizes the importance of consumption smoothing over the life-time (Sutch, 1991). Ransom and Sutch (1986, 1995) and Carter and Sutch (1996) argue that ‘modern’ retirement was much more widespread among workers before the passage of Social Security legislation than is generally accepted.[23] Planned retirement in turn required asset accumulation during prime working age (Carter and Sutch, 1996, pp. 5-6). The well-known rise in the aggregate U.S. saving rate over the nineteenth century is invoked as evidence of the increased importance of life-cycle motives, but Sutch and his collaborators present no micro-level evidence on saving behavior to support their argument.

The weight of microeconomic evidence does not support this interpretation. Wealth accumulation by most workers was quite modest by the end of middle age. James, Palumbo, and Thomas (2005) simulate lifetime family wealth accumulation in a model in which the underlying dynamics are estimated from our saving data set using Moffitt’s (1993) method of estimating dynamic relationships based on a set of repeated cross-sections. Median families in the simulation accumulated only about three years worth of income in net worth by age 55, about half of which was in the form of housing wealth. Families below the median accumulated much less financial wealth, often falling under an average year’s worth of income. Further evidence suggests that the modest wealthholdings by the elderly were generally supplemented by private intergenerational transfers from children to parents (reverse bequests) in a *de facto* pay-as-you-go system similar to later state-run social insurance schemes.

[23] For criticism, see Margo (1993).

The microeconomic evidence on the age profile of savings by working households also runs counter to the traditional life-cycle model. In the basic model, abstracting from uncertainty about future labor earnings (or other relevant economic variables), forward-looking agents use capital markets, borrowing and their own saving to attempt to maintain constant marginal utility of expenditure over time. From this follows the well-known implication that the shape of the lifetime consumption path should be independent of the shape of the path of expected earnings. In principle, a straight-forward method of examining the accuracy of the life-cycle hypothesis using survey data would be to compare age profiles of income and consumption (Carroll and Summers, 1991). What do our data show?

Our data set consists only of working respondents, so at higher ages we should not necessarily expect to observe substantial dissaving (though some of these older respondents might have been partially 'retired'). However, at younger ages we would certainly expect to see a pattern of 'hump saving' if respondents anticipated a period of low earnings in old age. Fig. 4 shows five-year moving averages of average consumption and income by age of the respondent, classified by skill category—unskilled, semiskilled, skilled, or white collar/professional—to allow for differently-shaped lifetime income profiles. Panel A shows the results based on reported saving with consumption measured as the difference between income and saving; panel B shows the results for households whose consumption is reported directly, but whose saving is calculated as the difference between spending and income. Consumption tracks income very closely in all skill groups—not what we would expect from the stylized life-cycle model. This result should not be too much of a surprise. Similar parallel relationships have been found for a number of other time periods and countries—e.g., for various occupational and educational groups in the contemporary U.S. (Carroll and Summers, 1991, pp. 318-27) and in less-developed countries such as the Ivory Coast and Thailand (Deaton, 1997, pp. 339-40).[24]

The close tracking of savings and income might be explained by impatient households faced by liquidity constraints (Thurow, 1969)—households would like to spend more, but cannot borrow. If savings were governed by a combination of impatience and liquidity constraints, we would expect consumption to exhaust income, as households continually press against their

[24] To be sure, this demonstration is not iron-clad. The profiles here are just cross sections of independent households, not evidence on the behavior of particular households over time. Households at different ages may have had different lifetime experiences of earnings and wealth accumulation, which may have affected the income/consumption age profiles. Since our surveys come from a relatively small range of years, it is not easy to distinguish the impact of potential cohort effects.

budget constraints.[25] But only at young ages of the household head (for reported saving under age 25) do we find a coincidence of consumption and income. Among older households, the average savings rate is consistently positive, inconsistent with the suggestion that consumption for a majority of families was always limited by current income. The empirical predictions of constrained consumption are observationally equivalent to ‘hand-to-mouth’ behavior, which we have already rejected econometrically.[26]

It would be possible to interpret the absence of life-cycle behavior by American working households at the end of the nineteenth century as simply a reflection of the absence of life-cycle motives. But we are not disposed to make that case here. Although life expectancy at birth was surely lower a century ago, most of the difference was due to lower infant and child mortality. The evidence from Massachusetts in 1890 indicates that the average 20 year-old male could expect to live another 40.7 years; the average 40 year-old, another 27.4 years; the average 60 year-old, another 14.7 years.[27] Lengthy lives alone, however, are inconclusive. After all, life-cycle behavior involves saving for *planned* retirement and workers may not have expected to leave the workforce. Certainly, the traditional account of retirement, in which workers were presumed unable to afford retirement before the introduction of social security in 1935, is consistent with this argument. Lee’s analysis of the 1900 Census (1996), in which he finds that less than 5 % of workers over 55 retired, supports this general interpretation, contra Sutch. Most workers did not save for retirement, not because they anticipated an early demise, but because they never expected to retire; an extended expectation of life was accompanied by an extended expectation of work. Older workers were, however, more susceptible to income loss from nonemployment (via sickness, as well as job loss) and were therefore not entirely immune from age-specific considerations in making saving decisions.[28] The evidence on family structures,

---

[25] The additional assumption that produces parallel but not coincident profiles of income and consumption over time is income uncertainty. See the discussion of precautionary savings below.

[26] It could be that the hump shape in family consumption was driven by changes in family size, as children are born at younger ages of the head and then later leave the household at older ages. While we do not have detailed information of the age structure of survey families, redoing Fig. 4 in per capita (of the family) terms still shows that consumption tracks income over the lifetime (not shown here to prevent a proliferation of figures, but available from the authors upon request).

[27] Note also the additional demands on wealth of lengthier female lives in combination with lower female labor force participation and age-gaps at marriage.

[28] The number of days lost by skilled manufactured workers rose from 29.5 on average for those aged 40-45 to 38.8 for those aged 60-65.

however, suggests that such shortfalls were normally covered by the extended family system in which elderly workers cohabited with their working children, rather than by prior saving (Thomas and Johnson, 2003).

5.2 Smoothing over the medium-term

Theories of precautionary saving emphasize the need for households to protect themselves against uncertain events, such as future income shocks, by building up 'a reserve against unforeseen contingencies' (Keynes, 1936, p. 107). Recent models of precautionary behavior that operate within the framework of intertemporal optimization (e.g., Carroll, 1997, 2001a, 2001b, Deaton, 1991, Hubbard, Skinner, and Zeldes, 1994, 1995) have been shown to capture the empirical regularities in contemporary (U.S.) data that are incompatible with the certainty equivalence model, including the parallelism of consumption and income, the non-zero marginal propensity to consume out of windfall income, and the relatively small holdings of wealth by households on the eve of retirement.

Can these models improve our understanding of the dynamics of working-class saving at the end of the nineteenth century? For a formal test of the importance of precautionary motives for saving in our data set, we follow Skinner (1988) in taking measured saving as a linear function of measured income, y, income uncertainty, ω, and demographic variables, Z.[29] To avoid estimation problems from measurement error in actual income, $y_{i\,(t),}$ we use a standard technique of grouping individual observations (see, e.g., Meghir, 2004, pp. F295-F296).[30] Thus,

$$S_{j\,(t)} = \beta_1 y_{j\,(t)} + \beta_2 \omega_j + \beta_3 Z_{j\,(t)} + \beta_0 + \nu_{j\,(t)} \qquad (2)$$

---

[29] In what follows, our focus is on medium frequency consumption smoothing (i.e. from year-to-year or over the business cycle). Annual data preclude examining within-year (high frequency) saving patterns directly, but elsewhere (James, Palumbo, and Thomas, 2001) we apply our historical savings data set to simulate high frequency saving for hypothetical families in a rational expectations model. We find that such smoothing could account for only about 28 % of actual savings levels. Thus, while working-class families may have used their saving to buffer earnings shocks within a given year, such motives still leave the bulk of saving yet to be accounted for.

[30] In estimating equation (1) such bias would have affected the estimate of the coefficient on unexplained income. With calculated saving as the dependent variable, the estimate of α_3 will be biased upward toward one, but estimates of the marginal propensity to save out of permanent and transitory income, α_1 and α_2, should be unaffected (Paxson, 1992, p. 19).

The group-level regressions specify cell-level, average saving as the dependent variable. Our group-level analysis, therefore, focuses on the calculated-saving subsample. Simply averaging the heavily left-censored reported-saving variable across families is inappropriate.

where observations are the means of groups j defined by age, skill, industry, and state across surveys.[31] Current income is measured in logs; demographic variables include age, marital status, and the number of dependents.

The specification is not ideal. More recent empirical work with precautionary saving models has typically focused on wealth as the dependent variable rather than saving (e.g., Carroll and Samwick, 1997, 1998; Lusardi, 1998; Cagetti, 2003). Carroll and Samwick (1997, pp. 43) observe, for example, 'As a theoretical proposition, the appropriate response to greater uncertainty is to hold more wealth.' We cannot measure wealth directly for most of the observations in our data set, nor do we have full information on saving as a net change in wealth-holdings. However, in a precautionary model, greater income volatility will generate a higher level of saving across all members of the group, as they try to establish a larger emergency fund, as well as more saving in the wake of a negative income shock as households attempt to replenish their reserves. As Carroll (1997, p. 39) notes 'groups of consumers with a greater variance of transitory shocks should on average exhibit a greater divergence between consumption and income.'

Secondly, in the absence of true panel data (à la Carroll and Samwick, 1998), we take as our measure of income uncertainty the intragroup standard deviation of log income where groups are defined by age, state, skill, and industry (instead of Skinner's occupational categories), a measure including both cross-sectional and time-series components.[32] To be sure, Carroll and Samwick (1997) have been very critical of Dardanoni's (1991) use of the within-group cross-sectional dispersion of income as an estimate for the expected variance of lifetime income. But the pre-World War I labor market was very different from today's. Income-experience profiles were much flatter, firm-specific capital less important, and longer-term relationships between firms and workers relatively rare (Jacoby, 1985; James, 1994; James and Thomas, 2002). Unemployment at this time has often been characterized as a lottery (Jacoby, 1985; Hatton and Williamson, 1991; James and Thomas, 2003). In this case, within-group dispersion would seem to have been a reasonable predictor of individual expectations of vulnerability to income shocks.

The results from OLS regressions on cell means are reported in Table 4. The baseline regression results appear in column (1); in column (2), the group-mean observations are

---

[31] As in Table 1, there are four age categories, four skill categories, and six industry categories.

[32] To ensure a reasonable component of time-series variability we exclude any groups on which there are not observations in at least three different years.

weighted by cellsize; in column (3), small cells with fewer than 44 observations are excluded, as in the IV regressions of Tables 2 and 3 (see note 22). The estimated coefficients on Logy are very stable and those for Sdlogy quite stable across specifications. More importantly for our purposes here, the estimated coefficients on our measure of income uncertainty, Sdlogy, are statistically significantly different from zero at conventional levels and of the expected positive sign in all alternative specifications.

Does statistical significance in this case also imply economic significance? In order to measure the quantitative important of precautionary saving, following Carroll and Samwick (1998), we calculate from the estimated equations in Table 4 what saving would have been if all groups had faced the minimum level of uncertainty for groups in the sample. The results for mean and median actual and simulated saving for specifications are shown in Table 5. If income uncertainty had been reduced to the minimum group level, mean and median saving would have fallen by between 55 and 75 %. These magnitudes moreover might well be understatements. If income is afflicted with classical measurement error, the observed variability of income will be greater than its true value, and the estimated coefficient will consequently be understated.[33] In addition, sampling error within the groups could introduce measurement error (Browning and Lusardi, 1996, p.1828), further biasing the estimated coefficient β_2 downward.

These results suggest that precautionary motives are a more likely explanation of the pattern of American savings at the end of the nineteenth century than life-cycle considerations. Our data show a higher level of precautionary saving than reported by Carroll and Samwick using similar methods for 1981-7. In section 1, we argued that the risks and/or uncertainties facing American workers in this period were substantial and that institutional mechanisms to deal with such shocks were relatively underdeveloped compared to the modern period. Prudential households will always save more, the greater the uncertainty they face (Hubbard, Skinner, Zeldes, 1994b, p. 63). Carroll's (2001a) simulations indicate that the higher the probability of a significant decline in income, the higher the target wealth-income ratio. It is likely that the probability of a catastrophic shortfall of income was much higher a century ago by virtue of the inherent volatility of the labor market (including the higher likelihood of disability or accident) and the paucity of workmen's compensation, unemployment or medical insurance schemes to replace lost incomes.

[33] We are indebted to an anonymous referee for this point.

Other empirical features of our data set are compatible with a precautionary model. The econometric results in Table 2 and 3 indicate marginal propensities to save out of permanent income that are statistically significantly greater than zero, and marginal propensities to save out of transitory income that are significantly less than one—values that are broadly consistent with the predictions of precautionary models.[34] The age-specific income-consumption profiles for the various occupational groups in Fig. 4 replicate the profiles for recent US households cited by Carroll (1997, pp. 34-8) as evidence against the traditional life-cycle model and in favor of precautionary motives, with one significant exception—Carroll's profiles show a growing decoupling of income and consumption in late middle-age (from 50 years old on), as life-cycle and retirement motives for saving become more powerful, while our Fig. 4, in contrast, shows that consumption tracks income for all age groups, consistent with the notion that retirement was an alien concept for most workers before 1914.[35]

Contemporary models of precautionary savings focus on risk-averse (prudent) households. In buffer-stock models (e.g., Deaton, 1991; Carroll, 1997), consumers prudently respond to uncertainty by accumulating an 'emergency reserve' of liquid assets; however, the size of the reserve is limited by the desire of households to consume today rather than in the future ('impatience'). The higher the rate of time preference relative to the real rate of interest, and the faster the expected growth rate of household income, the smaller the buffer stock; the higher the degree of uncertainty, and the higher the degree of risk aversion ('prudence'), the larger the buffer stock.[36] In the model of Hubbard, Skinner and Zeldes (1994, 1995), hereafter HSZ, households maximize their expected lifetime utility subject to 'uninsured idiosyncratic

---

[34] Carroll in simulations of buffer-stock models with a range of plausible parameters finds marginal propensities to consume (mpc) out of permanent income in the range of 0.8 to 0.95 (2001a), and mpc out of transitory income of around one-third (2001b). Our estimates of the mps out of permanent income are somewhat larger, and of the mps out of transitory income somewhat smaller. The deviations may reflect differences in the situation faced by late nineteenth century households relative to the assumptions underlying Carroll's simulations, in particular a higher degree of income uncertainty and more binding liquidity constraints. Our results are consistent with other empirical estimates, see note 22.

[35] Hubbard, Skinner and Zeldes (1994b, p. 86) note that their life-cycle model with uncertainty replicates the observed hump-shaped profile of consumption (in contrast to a life-cycle model with certainty). However, their simulations of earnings and consumption by age make it clear that their model does not replicate the close tracking of income and consumption observed in contemporary data (e.g. Carroll (1997, p. 34, Fig. IV), Browning and Crossley (2001, p. 13, Fig. 2)) or in our data.

[36] Deaton (1991) has argued that liquidity constraints further increase the size of the buffer stock, since households will be unable to borrow to cover unexpected shortfalls in income. Carroll (2001a), however, suggests that liquidity constraints are very much a secondary consideration in the determination of the size of the ideal buffer-stock; it is the combination of uncertainty, impatience and prudence that matters.

risk' (Hubbard, Skinner, and Zeldes, 1994a, p. 174), affecting incomes, lifespans and (medical) expenditures. Households are prudent, but not impatient (their rate of time preference being the same as the real rate of interest); they are, however, liquidity constrained.

It is not an easy task to determine which of the defining assumptions of these models (impatience, liquidity constraints) best fits our data. Rather than applying analytical solutions, the models employ numerical techniques calibrated to the specific circumstances of the economy they are designed to simulate (Hubbard, Skinner and Zeldes, 1994, p. 63; see also, Carroll, 2001b, pp. 27-9; Browning and Lusardi, 1996, p. 1807), thus reducing their generality to economies characterized by different institutional regimes.[37] Moreover, because of the computational-intensity of numerical solutions, sensitivity analysis of results is used sparingly. However, despite the need to be tentative when moving from models designed to shed light on a world a century removed from our data, we can draw some conclusions by focusing on a few of the more significant differences in the predictions of the two models for patterns of saving and wealth accumulation.

Both the buffer-stock and the HSZ models predict lower ratios of assets to income for the typical household than the certainty-equivalence model.[38] The differences between the two models focus on: i) the age profile of the wealth-income ratio (the buffer-stock model predicts a stable ratio during the years in which precautionary motives dominate saving behavior—roughly 25-50 in Carroll's specification (1997, pp. 42-5)— before rising as retirement becomes increasingly imminent, while the HSZ model predicts a rising ratio throughout the working life time); ii) the frequency distribution of wealth accumulation across the population (the HSZ model mimics the contemporary distribution very well, while the buffer-stock model underpredicts the wealth-income ratio in the upper reaches of the distribution[39]); and iii) the predicted number of households with very low savings rates in any given year (the buffer-stock

[37]For example, HSZ incorporate a consumption floor—a minimum level of consumption guaranteed by government transfers—which is clearly appropriate for the modern welfare state, but is irrelevant to the period we are studying.

[38] And therefore better fit contemporary US data, as evidenced by the PSID and other household surveys. Note that Carroll's model focuses on the ratio of financial (non-housing) wealth to permanent income of the median household, while the HSZ model is calibrated to the distribution of the ratio of total assets to income across the population.

[39] Carroll argues that this indicates either that there is a bequest motive at work, or that the population includes a certain proportion of the patient consumers, who build up a portfolio of liquid assets far in excess of that needed to maintain an emergency reserve. The HSZ model, in contrast, does not assume different motives for saving for different segments of the population.

model predicts a very small proportion, while the HSZ model predicts a much higher proportion of non-savers and low savers in the population. See Hubbard, Skinner, and Zeldes, 1994, pp. 87-9).

How do these different predictions relate to the historical evidence? We lack information on wealth-holding for most of the households in our data set. However, the Michigan Bureau of Labor Statistics asked workers in the agricultural implement and iron industries in 1890 for their present worth, the value of their home, and the amount still owed on the house. From the replies, we were able to calculate the value of non-housing and total wealth for over 7,800 households. The median ratio of non-housing wealth to (actual) income for 25-50 year olds was 0.72; a lower figure than would be predicted within a certainty-equivalence model, albeit higher than in recent US data (e.g., 0.29 in 1995). The median financial asset ratio rose from 0.67 for 30 year-olds to 0.86 for 50 year-olds and 1.04 for 65 year-olds; the median total asset ratio rose from 0.97 for 30 year-olds to 1.67 for 50 year-olds and 2.27 for 65 year-olds.[40] The distribution of wealth-holdings among the Michigan sample is surprisingly similar to modern survey data. The cumulative distribution function of the ratio of non-housing wealth to income in 1890 is almost identical to Carroll's reported cdf for 1995 (Carroll, 2001b, p. 40, Fig. 2); we have shown elsewhere (James, Palumbo and Thomas, 1995) that a simulation of asset accumulation using the entire data set generates asset ratios whose frequency distribution by age is remarkably similar to that depicted by HSZ from the PSID in the 1980s. Finally, we note that more than half the households in our overall data set were non-savers in any given year. Each of these findings suggest greater consistency in the historical data with the assumptions that underlie the HSZ rather than the (Carroll) buffer-stock version of precautionary behavior.

Late nineteenth-century working households appear then to have been risk-averse in the face of considerable income uncertainty, but did not behave in ways consistent with impatience. Was their behavior also shaped by strong liquidity constraints? The historical evidence is compatible with the HSZ model, which embraces liquidity constraints, but in the absence of a numerical solution with unconstrained households, it is not possible to determine how far their results depend on constraints as opposed to prudence and/or a minimum guaranteed income. Carroll (2001a) specifically addresses the issue of liquidity constraints in the context of a buffer-stock model; his simulations show that they are secondary to prudence in a model with

[40] In each case, these figures represent the wealth-income ratios for three-year averages centered on the reported age.

impatient consumers. However, since Carroll's model only applies sensitivity analysis to one dimension at a time, we cannot determine how far, for example, liquidity constraints matter less if households are prudent but not impatient. The separation of liquidity constraints from precautionary motives is not something that can be resolved econometrically: liquidity constrained households are unable to borrow, the precautionary motive induces 'self-imposed reluctance to borrow' for prudent households. If there is considerable uncertainty about future income, and a concern on the part of households that they may not be able to borrow to cover the totality of lost income, prudent households will tend to maintain a large emergency fund today, even if they are nowhere near their current budget constraint.

The evidence from our data set is likewise inconclusive. About six per cent of Michigan households who reported their present worth in 1890 declared that they were zero wealth-holders—a figure that is consonant with Carroll's simulation of households who are both impatient and constrained in a buffer-stock model, but which is rather lower than that predicted by the HSZ (1995) model. Three Kansas surveys in the 1880s and 1890s specifically asked workers if they had gone into debt in the previous year—some 24 per cent of respondents indicated that they had, indicating the absence of absolute binding constraints.[41] Of course, answers such as these do not address the issue adequately, since we do not know how many households would have chosen to have gone into debt had it been possible, nor do we know the levels of debt involved.

Our overall conclusion is that the historical evidence is consistent with a model in which households were motivated strongly by precautionary motives. As with the buffer-stock model, these households held much smaller stocks of assets than would be expected if they were pure life-cycle savers. However, the empirical evidence deviates in important ways from a buffer-stock interpretation. On the other hand, the success of the HSZ model in reproducing the levels and distribution of wealth-holdings by contemporary Americans largely turns on the assumption of a relatively high consumption floor (which reduces the incentive for households to save for retirement). The finding that their model predicts wealth-holdings for the late twentieth century that look remarkably similar to the accumulation levels of our households a century earlier thus creates something of a conundrum, since the institutional characteristics associated with the modern welfare state were far from the reality of the late nineteenth century.

[41] A similar question in a Maine survey in 1890 elicited positive responses from 20 per cent of respondents. See also fn. 2 above, which indicates that almost half of households in the Federal budget survey of 1902 either borrowed money or obtained credit to cover income shortfalls in the previous year.

However, if we consider the nature of the labor market and family economy a century ago, it becomes clear that historical circumstances also favored a low incentive to save for old age. The existence of family, co-residential arrangements, coupled with a highly uncertain economic environment and the absence of an expectation of retirement is crucial to understanding the weakness of life-cycle motives for saving and the low wealth-holdings typical of American working families at the end of the nineteenth century. This in turn adds another historical wrinkle—namely that any models of precautionary saving should include a measure of demographic uncertainty associated with the risks of outliving one's children as old-age providers.

6. Summary and conclusion

In this paper we examine the saving behavior of American workers around the turn of the century using a pooled set of independent cross-section workers' surveys. In the spirit of Paxson (1992), we estimate transitory income directly rather than treating it as a residual, using information on the variability of days lost (for reasons of illness and accident, as well as unemployment) and use this information to estimate marginal propensities to save out of permanent and transitory incomes in a pooled cross-section regression. We find that American workers around the turn of the century, a time before formal government social insurance programs existed, and before most workers had access to formal credit markets, were still on average able to smooth their consumption relative to income. Our econometric evidence consistently shows much larger estimated marginal propensities to save out of transitory income than out of permanent income, in alternative specifications and also allowing for possible unobserved heterogeneity, consistent with several other analyses of contemporary household saving behavior.

Although pre-World War I American households may have engaged in both high frequency and low frequency consumption smoothing, we find that most saving was motivated by medium frequency consumption smoothing, consistent with a precautionary model. If the degree of income uncertainty/dispersion facing households had been reduced to the lowest intragroup sample value, household saving would have declined by more than half. The dominance of medium frequency consumption smoothing is compatible with our understanding of the premodern labor market in the United States – a market in which individuals were faced with considerable riskiness via job loss, accident, sickness and premature death, promoting the need for precautionary saving, and in which institutional arrangements and cultural norms, such as lifetime working and extended families in which elderly workers resided with younger family members, limited the need for life-cycle saving. It should not be surprising therefore that the first moves towards mutual insurance, via labor unions and benefit societies, concentrated on accident, sickness, unemployment, strike pay and funeral coverage in their benefit schemes, and almost entirely ignored superannuation schemes.[42] Similarly, in the early push for social insurance by such agencies as the American Association for Labor Legislation (Rodgers, 1998, pp. 216-21, 241-66), the focus was on health insurance and workman's compensation rather than old-age pensions.

Appendix: Algebraic representation of the micro-level econometric model

This appendix presents the algebraic representation of the micro-level econometric saving model described in Sections 3 and 4. The complete model consists of three regression equations: one for predicting workdays lost (and then for estimating unexpected workdays lost) during the previous year; a second for explaining family income as a function of worker characteristics and unexpected workdays lost; finally, a third for estimating the marginal propensities to save out of permanent differences and transitory shocks to family income. The regression equations and fitted value definitions permit a simple algebraic representation:

$$DL_{i(t)} = X^{p}_{i(t)}{}'\,\gamma + \epsilon^{DL}_{i(t)} \qquad \text{(A1)}$$

$$\hat{DL}^{T}_{i(t)} = DL_{i(t)} - X^{p}_{i(t)}{}'\,\hat{\gamma} \qquad \text{(A2)}$$

$$Y_{i(t)} = X^{p}_{i(t)}{}'\,\beta^{P} + \hat{DL}^{T}_{i(t)}{}'\,\beta^{T} + X^{p}_{i(t)}\hat{DL}^{T}_{i(t)}{}'\,\beta^{I} + \epsilon^{Y}_{i(t)} \qquad \text{(A3)}$$

$$\hat{Y}^{P}_{i(t)} = X^{p}_{i(t)}{}'\,\hat{\beta}^{P} \qquad \text{(A4)}$$

$$\hat{Y}^{T}_{i(t)} = \hat{DL}^{T}_{i(t)}{}'\,\beta^{T} + X^{p}_{i(t)}\hat{DL}^{T}_{i(t)}{}'\,\beta^{I} \qquad \text{(A5)}$$

$$\hat{Y}^{U}_{i(t)} = Y_{i(t)} - \hat{Y}^{P}_{i(t)} - \hat{Y}^{T}_{i(t)} \qquad \text{(A6)}$$

---

[42] The only unions in the United States offering superannuation benefits for older workers were two branches of British origin, which in combination served fewer than 3,500 out of 440,000 unionized workers in 1897 (Bemis, 1899).

$$S_{i(t)} = \alpha_0 + \alpha_1 \hat{Y}^P_{i(t)} + \alpha_2 \hat{Y}^T_{i(t)} + \alpha_3 \hat{Y}^U_{i(t)} + \epsilon^S_{i(t)} \tag{A7}$$

$DL_{i(t)}$ stands for the number of workdays lost during the previous year reported by family i which was surveyed during year t. $X^P_{i(t)}$ is a vector of interacted indicator variables describing the worker's characteristics. Then $\hat{DL}^T_{i(t)}$ denotes our estimate of the unexpected component of the family's workdays lost realization. Similarly, $Y_{i(t)}$ is the family's reported income realization; $\hat{Y}^T_{i(t)}$ denotes our estimate of the transitory component to income, while $\hat{Y}^P_{i(t)}$ is our estimate of the permanent component of realized income. Finally, the residual, $\hat{Y}^U_{i(t)}$, measures the part of realized income which cannot be attributed either to the permanent or to the transitory component based on the regression equations.

The three equations of the complete model—(A1), (A3), and (A7)—are estimated sequentially. First, we estimate equation (A1) to obtain coefficient estimates to predict workdays lost for each family and then take the residual as unexpected workdays lost (A2). Second, we estimate regression (A3) using the first-stage residual to measure unexpected workdays lost in the income equation. We then use the second-stage income coefficients to estimate permanent income (A4), transitory income (A5), and unexplained income (A6, the residual from (A3)), all of which in turn are used to estimate the saving regression equation (A7). Bootstrap methods are used to estimate consistently standard errors for the coefficients in the third-stage saving regression (A7).

References

Alessie, Rob and Annamaria Lusardi (1997). 'Saving and Income Smoothing: Evidence from Panel Data', *European Economic Review*, **41**, 1251-79.

Bemis, Edward W. (1899). 'Benefit Features of American Trade Unions', *Bulletin of the Department of Labor*, **IV**, 361-400.

Browning, Martin and Thomas F. Crowley (2001). 'The Life-Cycle Model of Consumption and Saving', *Journal of Economic Perspectives*, **15**, 3-22.

Browning, Martin and Annamaria Lusardi (1996). 'Household Saving: Micro Theories and Micro Facts', *Journal of Economic Literature*, **34**, 1797-1857.

Cagetti, Marco (2003). 'Wealth Accumulation over the Life Cycle and Precautionary Savings', *Journal of Business and Economic Statistics*, **21**, 339-53.

Carroll, Christopher D. (2001a). 'Precautionary Saving and the Marginal Propensity to Consume out of Permanent Income', NBER Working Paper 8233.

Carroll, Christopher D. (2001b). 'A Theory of the Consumption Function, With and Without Liquidity Constraints', *Journal of Economic Perspectives*, **15**, 23-45.

Carroll, Christopher D. (1997). 'Buffer-Stock Saving and the Life-Cycle/Permanent Income Hypothesis', *Quarterly Journal of Economics*, **112**, 1-55.

Carroll, Christopher D. and Andrew Samwick (1998). 'How Important is Precautionary Saving?', *Review of Economics and Statistics*, **80**, 410-19.

Carroll, Christopher D. and Andrew Samwick (1997). 'The Nature of Precautionary Wealth', *Journal of Monetary Economics*, **40**, 41-71.

Carroll, Christopher D. and Lawrence H. Summers (1991). 'Consumption Growth Parallels Income Growth: Some New Evidence', in B.D. Bernheim and J.D. Shoven (eds), *National Saving and Economic Performance*, University of Chicago Press, Chicago, 305-43.

Carter, Susan and Richard Sutch (1996). 'Myth of the Industrial Scrap Heap: A Revisionist View of Turn-of-the-Century American Retirement', *Journal of Economic History*, **56**, 5-38.

Dardanoni, Valentino (1991). 'Precautionary Savings under Income Uncertainty: A Cross-Sectional Analysis', *Applied Economics*, **23**, 153-60.

Deaton, Angus (1997). *The Analysis of Household Surveys*, The Johns Hopkins Press, Baltimore.

Deaton, Angus (1992). *Understanding Consumption*, Oxford University Press, Oxford.

Deaton, Angus (1991). 'Saving and Liquidity Constraints', *Econometrica*, **59**, 1221-48.

Deaton, Angus (1990). 'Saving in Developing Countries: Theory and Review', in *Proceedings of the World Bank Annual Conference on Development Economics*, 1989, World Bank, Washington, D.C., 61-96.

De Long, J. Bradford and Lawrence H. Summers (1986). 'The Changing Cyclical Volatility of Economic Activity in the United States', in R.J. Gordon (ed.), *The American Business Cycle*, University of Chicago Press, Chicago, 679-719.

Engen, Eric and Jonathan Gruber (2001). 'Unemployment Insurance and Precautionary Saving', *Journal of Monetary Economics*, **47**, 545-79.

Flavin, Marjorie (1991). 'The Joint Consumption/Assets Demand Decision: A Case Study in Robust Estimation', NBER Working Paper No. 3802.

Hanes, Christopher (1993). 'The Development of Nominal Wage Rigidity in the Late 19th Century', *American Economic Review,* **83**, 732-56

Hatton, T.J. and J.G. Williamson (1991). 'Unemployment, Employment Contracts, and Compensating Wage Differentials: Michigan in the 1890s', *Journal of Economic History*, **51**, 605-32.

Hubbard, R. Glenn, Jonathan Skinner, and Stephen Zeldes (1995). 'Precautionary Saving and Social Insurance', *Journal of Political Economy*, **103**, 360-99.

Hubbard, R. Glenn, Jonathan Skinner, and Stephen Zeldes (1994a). 'Expanding the Life-cycle Model: Precautionary Saving and Public Policy', *American Economic Review Papers and Proceedings*, **84**, 174-179.

Hubbard, R. Glenn, Jonathan Skinner, and Stephen Zeldes (1994b). 'The Importance of Precautionary Motives in Explaining Individual and Aggregate Saving', *Carnegie-Rochester Conference Series on Public Policy*, **40**, 59-126.

Industrial Commission (1901). *Report of the Industrial Commission on the Relations and Conditions of Capital and Labor,* Government Printing Office, Washington, D.C.

Jacoby, Sanford (1985). *Employing Bureaucracy*, Columbia University Press, New York.

James, John A. (1995). ‘Reconstructing the Pattern of American Unemployment before World War I’, *Economica*, **62**, 291-311.

James, John A. (1994). ‘Job Tenure in the Gilded Age,’ in G. Grantham and M. MacKinnon (eds), *Labour Market Evolution*, Routledge, London, 185-204.

James, John A. (1993). ‘Changes in Economic Instability in 19th-Century America’, *American Economic Review*, **83**, 710-31.

James, John A. and Mark Thomas (2004). ‘Romer Revisited: Long-term Changes in the Cyclical Sensitivity of Unemployment’, unpublished manuscript, University of Virginia.

James, John A. and Mark Thomas (2003). ‘A Golden Age? Unemployment and the American Labor Market, 1880-1910’, *Journal of Economic History*, **63**, 959-94.

James, John A. and Mark Thomas (2002). ‘Labor Hoarding, Skills, and Productivity during Downturns in the Pre-Modern Labor Market: Evidence from the Late Nineteenth-Century United States’, unpublished manuscript, University of Virginia.

James, John A., Michael Palumbo, and Mark Thomas (2005). ‘Have American Workers Always Been Low Savers? Patterns of Accumulation among Working Households, 1885-1910’, *Research in Economic History*, **23**, 127-76..

James, John A., Michael Palumbo, and Mark Thomas (1998). ‘”What’s the Frequency, Kenneth?” Consumption Smoothing at High- and Medium- Frequencies among Working-Class Americans around the Turn of the Century’, unpublished manuscript, University of Virginia.

Johnson, Paul (1985). *Saving and Spending: the Working-class Economy in Britain, 1870-1914*, Clarendon Press, Oxford.

Kantor, Shawn and Price Fishback (1996). ‘Precautionary Saving, Insurance, and the Origins of Workers’ Compensation’, *Journal of Political Economy*, **104**, 419-42.

Keyssar, Alexander (1986). *Out of Work*, Cambridge University Press, Cambridge.

Lusardi, Annamaria (1998). ‘On the Importance of the Precautionary Saving Motive’, *American Economic Review Papers and Proceedings*, **88**, 449-53.

Margo, Robert. A. (1993). ‘The Labor Force Participation of Older Americans in 1900: Further Results’, *Explorations in Economic History*, **30**, 409-23.

Meghir, Costas (2004). ‘A Retrospective on Friedman’s Theory of Permanent Income’, *Economic Journal*, **114**, F293-F306.

Moffitt, Robert (1993). 'Identification and Estimation of Dynamic Models with a Time Series of Repeated Cross-Sections', *Journal of Econometrics* , **59**, 99-123.

Olney, Martha L. (1998). 'When Your Word is Not Enough: Race, Collateral, and Household Credit', *Journal of Economic History*, **58**, 408-31.

Palumbo, Michael G. (2000). 'Estimating the Effects of Earnings Uncertainty on Families' Saving and Insurance Decisions', *Southern Economic Journal*, **67**, 64-86.

Patterson, W. R. (1899). 'Pawnbroking in Europe and the United States', *Bulletin of the Department of Labor*, **IV**, 173-310.

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

Pistaferri, Luigi (2001). 'Superior Information, Income Shocks and the Permanent Income Hypothesis', *Review of Economics and Statistics*, **83**, 465-76.

Ransom, Roger and Richard Sutch (1995). 'The Impact of Aging on the Employment of Men in Working Class Communities at the End of the Nineteenth Century: A Cross-Section Analysis of Surveys from Maine, New Jersey, California, Michigan, and Kansas', in D. Kertzer and P. Laslett (eds), *Aging in the Past: Demography, Society, and Old Age*, University of California Press, Berkeley, 303-27.

Ransom, Roger and Richard Sutch (1986). 'The Labor of Older Americans: Retirement of Men on and off the Job, 1870-1937', *Journal of Economic History*, **46**, 1-30.

Romer, Christina (1986). 'Spurious Volatility in Historical Unemployment Data', *Journal of Political Economy*, **94**, 1-37.

Rotella, Elyce and George Alter (1993). 'Working Class Debt in the late Nineteenth Century United States', *Journal of Family History*, **18**, 111-34.

Skinner, Jonathan (1988). 'Risky Income, Life Cycle Consumption, and Precautionary Savings', *Journal of Monetary Economics*, **22**, 237-55.

Sundstrom, William A. (1990). 'Was There a Golden Age of Flexible Wages? Evidence from Ohio Manufacturing, 1892-1910', *Journal of Economic History*, **50**, 309-20.

Thomas, Mark and Paul Johnson (2003). 'Paying for Old Age: Past, Present, Future', in Paul A David and Mark Thomas (eds), *The Economic Future in Historical Perspective*, Oxford University Press for the British Academy, Oxford, 479-508.

Townsend, Robert M. (1995). 'Consumption Insurance: An Evaluation of Risk-Bearing Systems in Low-Income Economies', *Journal of Economic Perspectives*, **9**, 83-102.

U.S. Commissioner of Labor (1904). *Eighteenth Annual Report, 1903*. Government Printing Office, Washington, D.C.

Weir, David (1992). 'A Century of U.S. Unemployment, 1890-1990: Revised Estimates and Evidence for Stabilization', in R. Ransom (ed.), *Research in Economic History*, Vol. 14, JAI Press, Greenwich, Ct., 301-46.

Table 1
Summary of predicted and unexpected workdays lost and annual income during the previous year, by age, skill, and industry group

	Average workdays lost	Interquartile range for unexpected workdays lost 25%	75%	Average income	Interquartile range for unexpected annual income 25%	75%
I. Total sample	36.8	-28.3	17.4	$544.1	-31.5	54.2
II. By age group						
15-25 years	38.1	-31.3	16.4	424.0	-27.2	48.3
26-40 years	34.5	-26.9	17.4	594.5	-33.1	54.6
41-55 years	38.7	-28.8	18.7	606.2	-36.1	66.7
56-75 years	47.6	-34.8	24.2	538.2	-38.0	60.2
III. By skill class						
Unskilled	42.5	-32.1	21.4	412.5	-29.1	48.0
Semiskilled	38.2	-32.2.	19.1	511.5	-30.5	52.9
Skilled	36.5	-28.3	17.4	624.9	-35.6	70.9
White collar	27.7	-25.4	5.1	546.7	- 11.2	54.6
IV. By industry						
Manufacturing	37.0	-30.6	19.1	573.5	-36.6	66.3
Mining	62.2	-52.9	39.7	454.2	-57.2	81.9
Transportation	26.2	-25.0	5.1	605.4	-11.2	46.8
Construction	81.0	-46.4	32.6	583.7	-50.6	64.3
Trade	29.6	-29.5	9.6	595.5	-6.8	22.3
Services	19.9	-19.5	-3.5	662.0	6.7	40.1

Table 2
Estimates of equation (1) based on reported annual saving
(bootstrapped standard errors in parentheses)

I. Tobit estimates — **Alternative specifications**

Explanatory variable	**(1)**	**(2)**	**(3)**	**(4)**
Permanent income, $\hat{Y}^P$	0.3040 (0.0115)	0.3520 (0.0180)	0.3496 (0.0110)	0.2753 (0.0139)
Transitory income, $\hat{Y}^T$	0.5867 (0.0183)	0.4696 (0.0310)	0.5927 (0.0182)	0.5855 (0.0176)
Unexplained income, $\hat{Y}^U$	0.4194 (0.0093)	0.3271 (0.0160)	0.4207 (0.0091)	0.4164 (0.0098)
Age				3.276 (0.8047)
Age-squared				-0.0413 (0.0106)
Intercept	-205.705 (5.961)	-235.530 (6.145)	-227.656 (6.208)	-247.944 (11.693)
test $\alpha_1=\alpha_2$				
F(1,26067)= significance level	196.30 (0.00)	86.27 (0.00)		
F(1,25570)=			144.14 (0.00)	
F(1,26065)=				216.34 (0.00)
test $\alpha_2=1$				
F(1,26067)=	592.21 (0.00)	854.12 (0.00)		
F(1,25570)=			622.31 (0.00)	
F(1,26065)=				507.32 (0.00)

II. Instrumental variables Tobit estimates

	Instrument sets	
Explanatory variable	**Cohort, year**	**Age, skill, industry, state, year**
Permanent income, $\hat{Y}^P$	0.3009 (0.0179)	0.3549 (0.0124)
Transitory income, $\hat{Y}^T$	0.9305 (0.2138)	0.7801 (0.1262)
Intercept	-211.325 (10.168)	-240.578 (7.627)

Notes:

(1) contains baseline regression results based on all 26,070 reported saving observations.
(2) contains results based on sample of 26,070 in which all observations have been reweighted such that their sample contributes 10,000 members to the regression. This treatment has the effect of reducing the impact of Michigan surveys on the regression results.
(3) contains results based on a sample of 25,573 observations from Kansas and Michigan surveys only.
(4) contains results from a specification based on all 26,070 observations of reported saving with 'age' and 'age-squared' terms included in the regression equation.

Table 3
Estimates of equation (1) based on calculated annual saving
(bootstrapped standard errors in parentheses)

I. OLS estimates — **Alternative specifications**

Explanatory variable	**(1)**	**(2)**	**(3)**	**(4)**
Permanent income, $\hat{Y}^P$	0.3060 (0.0151)	0.3085 (0.0240)	0.2647 (0.0183)	0.3402 (0.0155)
Transitory income, $\hat{Y}^T$	0.4952 (0.0149)	0.5484 (0.0262)	0.4876 (0.0154)	0.4944 (0.0150)
Unexplained income, $\hat{Y}^U$	0.4155 (0.0123)	0.4095 (0.0198)	0.4095 (0.0144)	0.4161 (0.0122)
Age				-5.897 (0.8782)
Age-squared				0.0639 (0.0106)
Intercept	-75.667 (6.644)	-76.817 (5.960)	-52.884 (8.511)	28.317 (15.335)
test $\alpha_1=\alpha_2$				
F(1,12008)=	153.37	392.23		
significance level	(0.00)	(0.00)		
F(1,9350)=			116.45 (0.00)	
F(1,12006)=				105.29 (0.00)
test $\alpha_2=1$				
F(1,12008)=	1129.61 (0.00)	972.29 (0.00)		
F(1,9350)=			943.44 (0.00)	
F(1,12006)=				1141.49 (0.00)

II. Instrumental variables estimates

	Instrument sets	
Explanatory variable	**Cohort, year**	**Age, skill, state, industry, year**
Permanent income, $\hat{Y}^P$	0.2065 (0.0397)	0.2657 (0.0176)
Transitory income, $\hat{Y}^T$	0.9739 (0.1633)	1.0538 (0.1121)
Intercept	-18.506 (22.675)	-53.592 (9.738)

Notes:

(1) contains baseline regression results based on all 12,012 observations on calculated saving (reported income minus reported total expenditures).
(2) contains results based on sample of 12,012 in which all observations have been reweighted such that their sample contributes 10,000 members to the regression. This treatment has the effect of reducing the impact of Kansas surveys on the regression results.
(3) contains results based on a sample of 9,354 observations from Kansas and Michigan surveys only.
(4) contains results from a specification based on all 12,012 observations of reported saving with 'age' and 'age-squared' terms included in the regression equation.

Table 4
Estimates of equation (2) based on calculated annual saving
(standard errors in parentheses)

	Alternative specifications Grouped OLS		
Explanatory variable	**(1)**	**(2)**	**(3)**
Logy	220.730	217.243	198.964
	(29.179)	(26.718)	(35.220)
Sdlogy	262.135	196.383	279.481
	(81.436)	(85.375)	(92.806)
Age	-25.362	-32.975	-32.250
	(13.434)	(15.416)	(14.932)
Age-squared	0.299	0.385	0.351
	(0.160)	(0.187)	(0.174)
Married	124.480	173.637	227.17
	(58.628)	(61.653)	(68.351)
Dependents	-25.304	-16.521	-28.333
	(8.239)	(9.257)	(9.314)
Intercept	-892.157	-774.861	-657.093
	(238.251)	(270.56)	(273.448)
Adjusted R^2	0.4118	0.3990	0.4388

Notes:
(1) contains baseline OLS regression results based on 116 observations of grouped data with group-average calculated saving as the dependent variable.
(2) contains OLS regression results based on 116 observations of grouped data weighted by cellsize with group-average calculated saving as the dependent variable.
(3) contains OLS regression results on 89 groups with cellsizes of greater than 44; group-average calculated saving is the dependent variable.

Table 5
Actual and simulated saving levels

Specification		Mean	Median
(1)	Actual	$ 124.82	$ 124.66
	Simulated	30.59	36.27
(2)	Actual	111.84	109.53
	Simulated	41.26	39.01
(3)	Actual	92.67	91.61
	Simulated	41.78	40.30

Notes: For equation specifications see notes to Table 4.

Table A-1
Survey composition of the data set

	Surveys with reported saving		Surveys with calculated saving	
State	Year	No. of obs.	Year	No. of obs.
Maine			1886.5	62
			1887.5	87
			1888	88
			1890	1011
			1894	504
			1900	102
New Hampshire			1886	308
			1887	216
Michigan	1888	715		
	1889	4265	1889	1910
	1890	5922	1890	3801
	1892.5	5028		
	1895	3001		
	1896	3757		
Kansas			1884.5	349
	1885.5	316	1885.5	403
	1886.5	323	1886.5	393
	1895	233	1895	384
	1896	302	1896	426
			1899	724
	1903	617	1903	594
	1904	369	1904	368
	1905.5	329		
	1906.5	396		
Missouri	1890.75	255	1890.75	163
Oklahoma	1908	242	1909	117

Table A-2
Summary statistics for annual income and expenditure by age, skill and industry groups

I. Based on reported saving

	No. of obs.	Expenditure	Income	Var(C)/Var(Y)
A. Total sample	26,070	$481.44	$543.01	0.75
Nonsavers	15,254	475.27	475.27	
Savers	10,816	487.72	638.55	
B. By age group				
15-25 years	8,660	379.47	419.42	0.79
Nonsavers	5,711	371.40	371.40	
Savers	2,949	395.10	512.41	
26-40 years	12.232	531.55	606.85	0.77
Nonsavers	6,498	543.25	543.25	
Savers	5,734	518.29	678.93	
41-55 years	4,368	541.44	611.59	0.73
Nonsavers	2,575	536.91	536.91	
Savers	1.793	547.96	718.84	
56-75 years	810	459.14	530.52	0.72
Nonsavers	470	459.98	459.98	
Savers	340	457.99	628.03	
C. By skill class				
Unskilled	3,703	378.39	409.80	0.90
Nonsavers	2,513	387.17	387.17	
Savers	1,190	359.86	457.58	
Semiskilled	8,791	453.27	506.24	0.79
Nonsavers	5,451	446.16	446.16	
Savers	3,340	464.87	604.31	
Skilled	9,654	541.90	628.23	0.75
Nonsavers	4,826	543.67	543.67	
Savers	4,828	540.14	712.76	
White collar	3,922	486.41	541.44	0.79
Nonsavers	2,464	495.59	495.59	
Savers	1,458	470.88	618.93	
D. By industry				
Manufacturing	15,348	444.53	507.52	0.73
Nonsavers	8,560	427.90	427.90	
Savers	6,788	465.50	607.92	

Mining	924	404.68	445.23	0.84
Nonsavers	614	412.63	412.63	
Savers	310	388.93	509.78	
Transportation	8,616	538.25	600.51	0.75
Nonsavers	5,456	538.75	538.75	
Savers	3,160	537.38	707.15	
Construction	768	570.48	645.16	0.91
Nonsavers	395	612.52	612.52	
Savers	373	525.97	679.73	
Trade	74	633.70	735.13	0.68
Nonsavers	36	696.59	696.59	
Savers	38	574.11	771.64	
Services	340	605.81	681.44	0.98
Nonsavers	193	659.25	659.25	
Savers	147	535.65	710.57	

II. Based on calculated saving

	No. of obs.	Expenditure	Income	Var(C)/Var(Y)
A. Total sample	12,012	$476.22	$579.42	0.70
Nonsavers	4,376	497.85	474.30	
Savers	7,636	463.83	639.66	
B. By age group				
15-25 years	1,545	399.50	498.06	0.88
Nonsavers	568	430.38	413.48	
Savers	977	381.55	547.23	
26-40 years	6,695	480.97	586.17	0.76
Nonsavers	2,331	502.67	480.42	
Savers	4,364	469.38	642.65	
41-55 years	3,104	509.47	611.96	0.64
Nonsavers	1,195	527.48	498.52	
Savers	1,909	498.20	682.97	
56-75 years	668	451.62	548.72	0.63
Nonsavers	282	468.39	443.56	
Savers	386	439.38	625.55	
C. By skill class				
Unskilled	1,643	371.60	420.97	1.00
Nonsavers	785	383.46	360.05	
Savers	858	360.75	476.71	

Semiskilled	4,367	451.24	536.55	0.77
Nonsavers	1,781	475.20	451.13	
Savers	2,586	434.75	595.37	
Skilled	5,451	526.72	653.40	0.69
Nonsavers	1,654	565.38	542.82	
Savers	3.797	509.88	701.57	
White collar	551	486.64	659.67	0.70
Nonsavers	156	616.14	587.09	
Savers	395	435.50	688.33	
D. By industry				
Manufacturing	8,564	466.08	568.65	0.71
Nonsavers	3,102	479.12	463.39	
Savers	5,462	458.67	628.43	
Mining	468	436.09	498.59	0.69
Nonsavers	180	449.50	381.82	
Savers	288	427.70	571.57	
Transportation	850	602.44	728.69	0.70
Nonsavers	315	638.49	589.11	
Savers	535	581.23	810.87	
Construction	1,445	474.1	558.54	0.71
Nonsavers	564	505.14	471.80	
Savers	881	454.23	614.09	
Trade	217	462.01	577.33	0.56
Nonsavers	73	515.85	485.65	
Savers	144	434.71	623.80	
Services	468	485.93	651.50	0.69
Nonsavers	142	618.21	579.25	
Savers	326	428.32	682.97	

Figure 1. Density and Distribution of Reported Saving

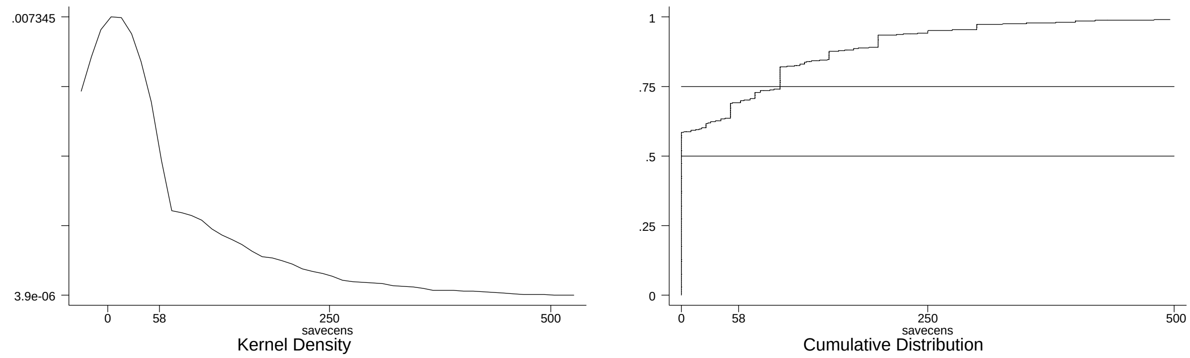


Figure 2. Density and Distribution of Calculated Saving

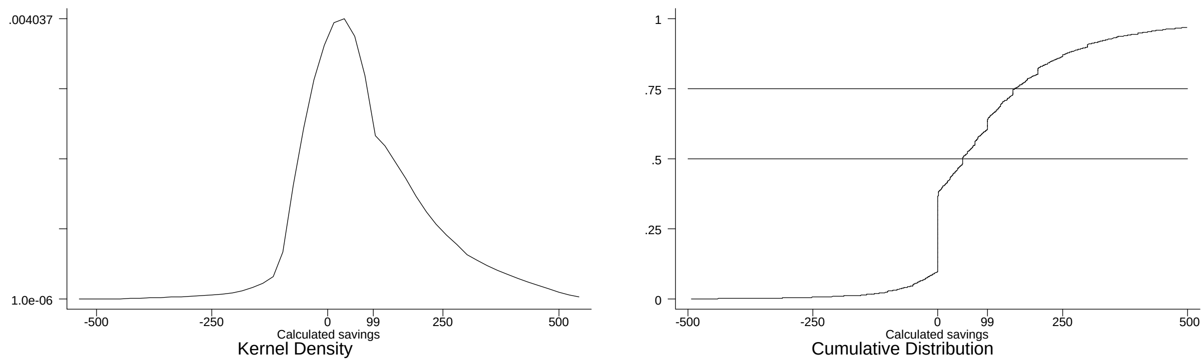


Figure 3. Density and Distribution of Unexpected Workdays Lost

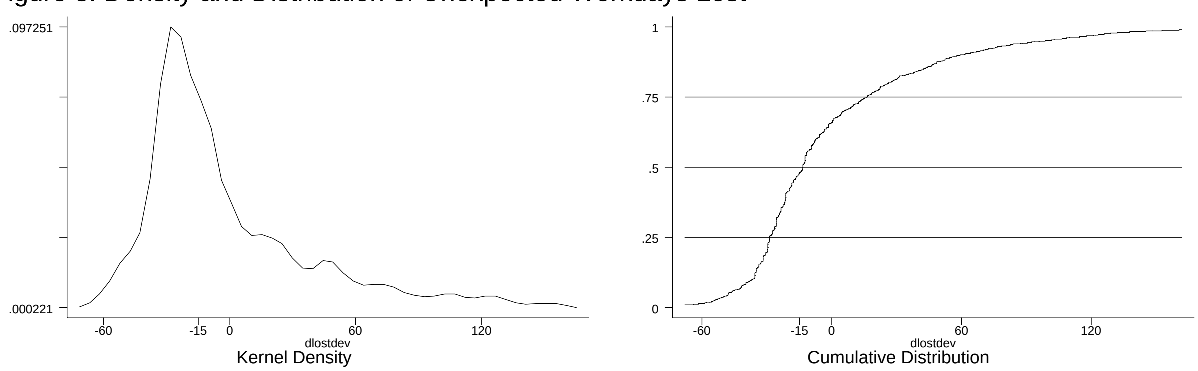
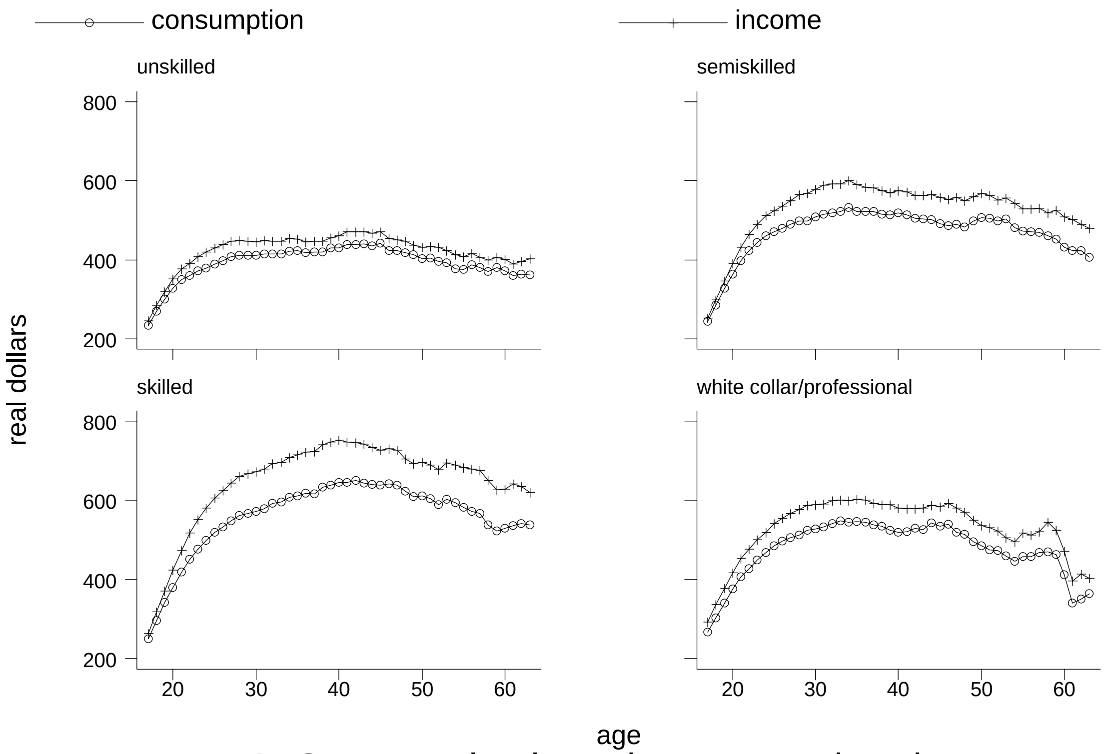
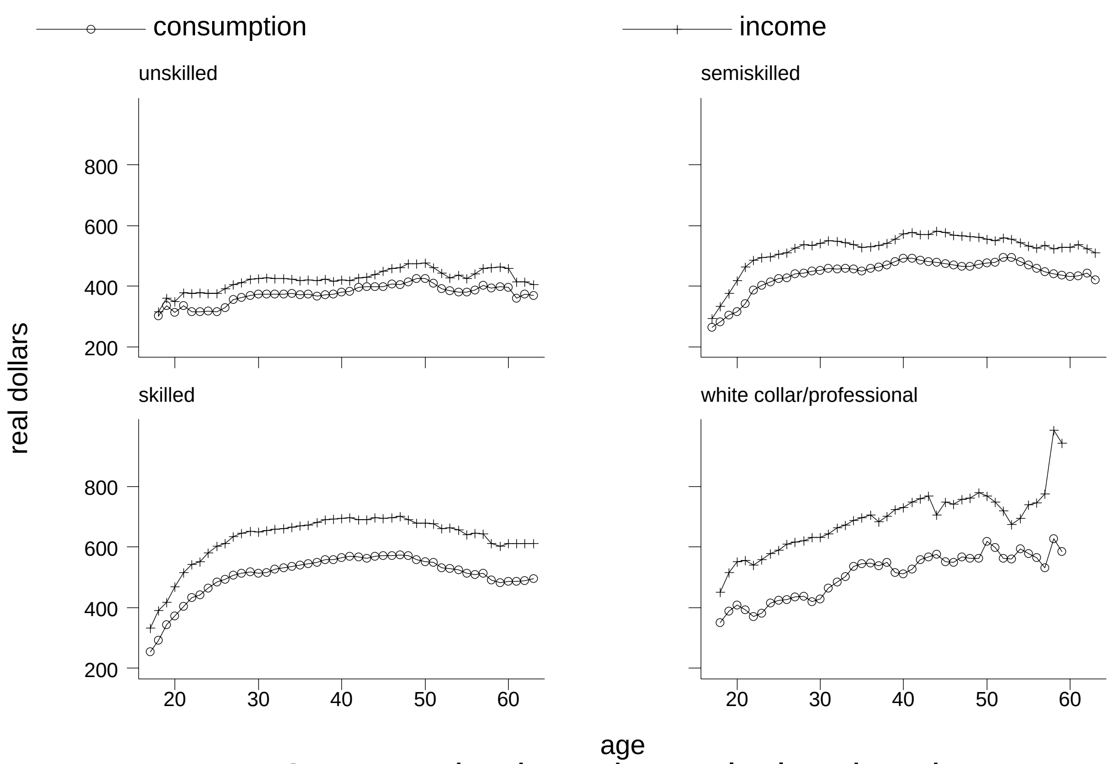


Figure 4. Consumption and Income Profiles by Skill Class



A. Consumption based on reported saving



B. Consumption based on calculated saving