

A Nonparametric Characterization of Income Uncertainty over the Lifecycle

James Feigenbaum* and Geng Li^{†‡}

July 14, 2008

Abstract

We propose a novel nonparametric approach to estimate household income uncertainty at various future horizons, and to characterize how income uncertainty evolves over the lifecycle. Our method acknowledges that households may have superior information about their future income than econometricians and exploits the intuition that if an income process exhibits more uncertainty then future income will be more difficult to forecast. Accordingly, we measure income uncertainty as the variance of forecast residuals. Our estimates generally imply smaller and less persistent income uncertainties than the previous literature has documented. In addition, we show that income uncertainty exhibits a U-shaped profile over the lifecycle, with young and old households facing greater uncertainty than middle-aged households. This pattern is particularly pronounced for near and intermediate forecast horizons. Our results are robust to various sample and model specifications and hold with respect to both labor and total family income. As opposed to a unit-root process, we find that our nonparametric estimates are better replicated by an income process with highly persistent, autoregressive shocks that have age and horizon dependent variances.

JEL Classification: D12, D91, E24

Keywords: income uncertainty, forecast errors, nonparametric estimates, lifecycle

*Department of Economics; University of Pittsburgh; 4906 W. W. Posvar Hall; 230 South Bouquet St.; Pittsburgh, PA 15260. E-mail: jfeigen@pitt.edu. URL: www.pitt.edu/~jfeigen.

[†]Federal Reserve Board. Email address: geng.li@frb.gov. The views presented in this paper are those of the author's and are not necessarily those of the Federal Reserve Board and its staff.

[‡]We thank Juan Carlos Cordoba, David DeJong, Karen Dynan, Jonathan Heathcote, Michael Palumbo, and seminar participants at the Conference on Computation in Economics and Finance, Federal Reserve Board, the Econometric Society North American Summer Meetings, SUNY Stony Brook, the University of Pittsburgh, and the University of Texas-Dallas for helpful discussions and comments. We also thank Meredith Richman for assistance in creating the charts. All remaining errors are our own.

1 Introduction

Income uncertainty plays an important role in household decisions regarding consumption, saving, and investment. First, in the presence of incomplete financial markets, households facing uninsurable income risk have to save for precautionary reasons (Aiyagari (1994)). How much they need to save crucially depends on the level of future income uncertainty. Second, people can choose their investment portfolios to achieve the optimal exposure to risks. How much people should invest in risky assets, such as stocks, also depends on their exposure to income risks (Viceira (2001)). Finally, choices of financial contracts and durable-good purchasing decisions are critically affected by household income uncertainty. For instance, Campbell and Cocco (2003) show that homeowners with more risky income should choose fixed-rate mortgages over adjustable-rate mortgages.

Many of these household financial and economic decisions exhibit strong lifecycle patterns. It is well known that nondurable good consumption has a hump-shaped lifecycle profile that closely tracks income (Gourinchas and Parker (2002)). Meanwhile, stock market participation rates are high among middle-aged households but puzzlingly low among seniors. Home ownership and durable goods consumption also tend to vary along the lifecycle (Fernández-Villaverde and Krueger (2007)).

Despite the pivotal importance of income uncertainty, the literature has not yielded a concrete measure of income uncertainty, and to the best of our knowledge even less work has been done to characterize how income uncertainty evolves over the lifecycle. This paper helps bridge this gap.

Previous researchers have typically taken the cross-sectional variance of income over a subpopulation to measure income variability. We argue that, while these measures of income variability are suitable for studying income inequality, they are not ideal for studying income uncertainty because these approaches assume that household expectations are based on an information set identical to what econometricians have. In contrast, Cunha, Heckman, and Navarro (2005) point out individuals may have additional, private information that is relevant to predicting their future income. Understating or exaggerating household private information can lead to over- or underestimation of the income uncertainty perceived by individuals, which will in turn distort our understanding of how income uncertainty influences consumption and saving behavior.

We propose a novel method of measuring income uncertainty that can be used to directly assess the significance of a household's superior information. Using panel data that provide information about a large number of households over many years, we estimate an income forecasting equation that includes additional information potentially available to households but not necessarily available to contemporaneous econometricians. Specifically, we assume that households take into account their impending demographic and working status changes when forecasting their future income. We interpret the sample variance of forecasting errors as the income uncertainty faced by households. We repeat

the forecasting exercise for various forecasting horizons and estimate income uncertainties associated with each horizon. Our approach therefore allows us to estimate a volatility matrix that summarizes the uncertainty households should expect as a function of both age and forecast horizon.

Another important innovation in our methodology is that we take a nonparametric approach in the sense that we do not specify any functional form for the data-generation process of household income. Rather, we estimate the second moments of forecasting errors at various horizons to quantify how variable and persistent innovations to the income process are. In comparison to the existing literature, because we incorporate household private information and do not rely on a unit-root assumption to identify the persistent component of income shocks, we find less persistence and a lower level of income uncertainty. Subsequently, we calibrate a flexible income process and we find that our nonparametric estimates can best be replicated by combining a persistent, but not unit-root, shock process and a transitory shock process, both with age-dependent variances.

The estimated volatility matrix also suggests that income uncertainty perceived by households evolves substantially over the lifecycle. For a given age in the future, we consistently find, as one would expect, that uncertainty about income at that age diminishes as the consumer approaches that age. For a fixed forecast horizon, when consumers are young, income uncertainty gradually declines with age, presumably as decisions on career, human capital development, and fertility are resolved. Income uncertainty reaches and stays at its lowest level during middle age. Afterwards, income uncertainty rises again, potentially due to uncertainty about working hours and health risks. Our U-shaped lifecycle profile of income uncertainty is consistent with Jaimovich and Siu (2008). Studying CPS data, they construct the business-cycle component of hours worked for workers of various age groups and find a U-shaped pattern in the volatility of hours worked by age.

We proceed as follows. Section 2 reviews the existing approaches of measuring income variability and sets forth the argument that these approaches are not ideal for measuring household income uncertainty. Section 3 introduces the empirical methodology that we employ in the current paper. Section 4 describes the data and sample construction. Section 5 presents the main results and shows that these results are robust to various modifications with respect to the model and the sample. Section 6 contrasts the current results with the income uncertainties reported in some influential earlier contributions. Section 7 calibrates a data-generating process to approximate our nonparametric estimates and to facilitate theoretical work. Section 8 concludes and sets an agenda for future work.

2 Common Practice and Some Critique

Traditionally, when rich micro data were not readily available, researchers had to infer the volatility of personal income from aggregate time-series data. The estimated volatility was often interpreted as the income uncertainty. More recent studies of aggregate output (see, among others, Kim and Nelson (1999), and McConnell and Perez-Quiros (2000)) have led to findings of the so-called “Great Moderation”, a sharp decline in the growth volatility of GDP and its components since the mid 1980s. Fogli and Perri (2006) attribute a substantial fraction of the decrease in personal savings and the rise of external balances to this reduction in aggregate volatility. However, aggregate data can mask important variations and correlations in income at the household level, so a decrease in the volatility of aggregate income does not necessarily imply that income at the household level has become less risky. To see this, consider a hypothetical economy populated by two consumers. Every period, each consumer receives one unit of endowment and then they engage in a zero-sum game of chance with their endowments. On account of this gambling, though the aggregate income of the economy remains constant and bears no uncertainty, individual households can expose their income to substantial risk.

Studying household income directly avoids this complication, and this has become feasible as large panel surveys of household income, such as the Panel Study of Income Dynamics (PSID), have become increasingly available. Previous work using household data has focused on the cross-sectional and/or time-series variances of household income.¹ For example, Dynan, Elmendorf and Sichel (2007) simply focus on the standard deviation across households of the percent change in household income.² More elaborate models often postulate some parsimoniously parameterized income process that includes a predictable part and a stochastic part, which in turn is some combination of permanent and transitory shocks. In practice, econometricians estimate the predictable part of income using a function of an age polynomial and other demographic variables as well as year dummies to filter out economy-wide variations (Carroll and Samwick (1997), Gourinchas and Parker (2002)). The residual, unexplained portion of income is then interpreted as the stochastic component. Various techniques can be applied to investigate the variance and persistence of this stochastic part to quantify income riskiness.

As an alternative approach, instead of estimating an income trend, some researchers take the mean income of a household over a given period as a proxy for permanent income and treat the difference between the observed income and the mean as the transitory component of income. In this manner, Gottschalk

¹To be fair, some of the methods we review here are not designed for studying income uncertainty and riskiness. Some approaches were developed to study income inequality dynamics, and some were developed to study income volatility in a statistical context. The caveats we raise mainly concern what we should be mindful about if these measures are interpreted as income uncertainties.

²Dynan, Elmendorf and Sichel (2007) also provide an elegant and comprehensive survey of the literature studying household income volatilities using household data.

and Moffitt (1994) argue that an increase in the variance of transitory earnings could explain a large portion of the widening of income inequality during the 1970s and 80s.

Essentially, what these methods attempt to measure is either the cross-sectional or the time-series variability of a specific component of income that is orthogonal to the specified information sets, conditional on which the predictable component of income is constructed. Conceptually, it is true that greater income uncertainty will create larger observed cross-sectional and/or time-series variability. However, it is not generically true that larger variability always implies greater uncertainty. We point out several important caveats that should be kept in mind when these approaches are used to study the underlying income uncertainty.

First, to accurately account for the income uncertainty faced by households, the predictable income component should be constructed conditional on the household's information instead of the information possessed contemporaneously by the econometrician, which may only be a small subset of the household's complete information. Should we condition on an information set more extensive than what the econometrician has, some part of the econometrician's residual term could be predicted. Indeed, Cunha, Heckman, and Navarro (2005) highlight the distinction between income heterogeneity (predictable variability) and uncertainty (unpredictable variability). They document that a substantial fraction (about 60 %) of returns to schooling is predictable at the time students decide to go to college. The unconditional variance of deviations from the sample mean, as in Gottschalk and Moffitt (1994), can exaggerate income uncertainty because this does not filter out the variance predictable to households.

As an illustrative example, consider the following simple hypothetical economy where household i 's income at time t is given by

$$y_{i,t} = e^{\psi_i t},$$

where ψ_i , the rate of income growth, is a random number assigned by nature to household i . Suppose ψ_i is known to the household but is not observed by the econometrician. If an econometrician measures income uncertainty as the dispersion of income relative to the mean of $y_{i,t}$ over the population, he will find that income uncertainty grows with respect to time, but this is a spurious conclusion since all households have a deterministic income process. Likewise, if the econometrician cannot distinguish the deterministic dispersion of income trends amongst households, the estimated time series volatility could also exaggerate the underlying income uncertainty.

Conversely, if the predictable income component is constructed conditional on information that households may not have, the estimated income uncertainty will contain a downward bias. In particular, households face income uncertainty with respect to income of year t only in years prior to t . Therefore, year- t income uncertainty is perceived conditional only on the information available to households before t .

As an example of how this point is relevant, consider the approach of

Carroll and Samwick (1997) and Gourinchas and Parker (2002) as an example.³ The logarithm of income, y_t , is decomposed according to

$$y_t = p_t + \varepsilon_t, \quad (1)$$

where ε_t is a transitory shock and p_t is a persistent shock. The latter is further assumed to obey

$$p_t = g_t + \rho p_{t-1} + \eta_t, \quad (2)$$

where g_t is predictable income growth, $\rho \in [-1, 1]$ is the autocorrelation of the persistent shocks, and η_t is the persistent income shock. When g_t is estimated, it is fitted conditional on time- t household demographic, occupation, and industry information. This is equivalent to assuming that at any time prior to t the households know all the time- t information useful to predict g_t , which could be an unnecessarily strong assumption.

Second, from a more technical point of view, many previous studies (e.g. Carroll and Samwick (1997) and Gourinchas and Parker (2002)) assume that the persistent income shocks are permanent, exhibiting the highest possible persistence. In terms of Eq. (2), ρ is assumed to be 1. As we show in Sections 6 and 7, the persistent shock is best modeled as a fairly persistent, but not a unit-root, process with ρ close to but strictly less than 1. This distinction will have a substantial effect on the level of income uncertainty over the lifecycle.

Third, fitting individual income or income growth with a trend only driven by age and other demographic characteristics requires assuming that all similar looking individuals more or less share a common life-cycle income trend. This assumption is consistent with the model introduced by MaCurdy (1982). A competing view is that individuals face individual-specific income profiles, as proposed by Lillard and Weiss (1979). Guvenen (2007) labels the MaCurdy-type income process as a restricted income profile, and the Lillard and Weiss-type income process as a heterogeneous income profile. He presents evidence that consumption data are more consistent with a heterogeneous income process. If the underlying income process is better characterized in this way, fitting it with a trend common to all households will increase the residual variance and hence exaggerate the underlying income uncertainty.

Finally, apart from a few exceptions, previous studies have not fully characterized how household income uncertainty evolves over the life cycle. It is both theoretically and empirically appealing to study whether the income uncertainty perceived by households does stay constant, and, if not, how it varies over the lifecycle. Intuitively, a single 22-year old college graduate first entering the labor market should have much greater uncertainty about his income five or ten years down the road than a 40-year old with a family and a settled career path should over the same time horizon.

³The empirical models estimating income uncertainty used in the two papers are almost the same. We use mostly the same notation as Carroll and Samwick (2002).

3 A Forecast-Based Non-Parametric Approach

This paper develops a nonparametric approach for measuring income uncertainty that is flexible enough to adapt to various assumptions about the household's information set and is also able to characterize the evolution of income uncertainty over the lifecycle. Our key insight is that, heuristically, greater income uncertainty should make future income more difficult to forecast. Consequently, we use the forecast accuracy as a measure of the underlying uncertainty. The larger the variances of forecast errors are the greater uncertainty a household has regarding their future income.

Let income at age t be y_t and the s -period-ahead income $y_{t,s}$. At age t , the s -period-ahead income can be decomposed as

$$y_{t,s} = E[y_{t,s} | \mathcal{I}_t^H] + \varepsilon_{t,s}, \quad (3)$$

where $E[y_{t,s} | \mathcal{I}_t^H]$ is the mathematical expectation of age $t + s$ income, $y_{t,s}$, conditional on age- t information, $\mathcal{I}_t^H$, and $\varepsilon_{t,s}$ is an error term orthogonal to $\mathcal{I}_t^H$. We characterize lifecycle income uncertainties using a variance matrix, Ω , and a sequence of correlation matrices, Θ^ψ . The element $\omega_{t,s}$ of the Ω matrix is the variance of the s -year-ahead forecast errors of all age- t households,⁴ i.e.

$$\omega_{t,s} = \text{Var}[\varepsilon_{t,s}] = \text{Var}[y_{t,s} - E(y_{t,s} | \mathcal{I}_t^H)]. \quad (4)$$

Elements of the Θ^ψ matrix, $\theta_{t,s}^\psi$, are the correlation coefficients between the s -year-ahead forecast errors, $\varepsilon_{t,s}$, and the ψ -year-ahead forecast errors, $\varepsilon_{t,\psi}$, of age- t households, i.e.

$$\theta_{t,s}^\psi = \text{Corr}(\varepsilon_{t,s}, \varepsilon_{t,\psi}). \quad (5)$$

For example, elements of Θ^1 are the correlation coefficients between the s -year-ahead forecast errors and the 1-year-ahead forecast errors. This approach is nonparametric because it does not presume that income shocks follow any specific process. In examining the Ω and Θ^ψ matrices, we can directly study the lifecycle dynamics of income uncertainty.

One hurdle to implementing this strategy is that we do not know the joint distribution of $y_{t,s}$ and $\mathcal{I}_t^H$. As a result, we cannot compute $E[y_{t,s} | \mathcal{I}_t^H]$ directly. Indeed, we do not even know exactly what $\mathcal{I}_t^H$ encompasses. This is one of the main reasons why methods used in previous studies potentially yield inaccurate estimates of income uncertainty. To examine quantitatively the bias introduced by ignoring superior information possessed by households and to alleviate this bias, we experiment with two specifications. First, in what we label as the restricted information specification (RIS), we project $y_{t,s}$ solely conditional on $\mathcal{I}_t^R$, the information set that econometricians observe at the time when the household is at age t . $\mathcal{I}_t^R$ is the information that households *certainly* possess at age t . Second, in what we call the augmented information

⁴In our implementation, we take the age of the head of household as the age of the household.

specification (AIS), we project $y_{t,s}$ conditional on the augmented information set $\mathcal{I}_t^A$, where

$$\mathcal{I}_t^A = \mathcal{I}_t^R \cup \mathcal{I}_t^F. \quad (6)$$

The augmenting future information set, $\mathcal{I}_t^F$, contains elements that econometricians observe s years later, when the household becomes age $t + s$, the target age of the forecast. This is information that households *likely or possibly* know at age t . Put differently, $\mathcal{I}_t^F$ approximates the superior information that households possess but econometricians do not observe concurrently.

To fix the idea, we estimate the following RIS equation,

$$y_{i,t,s} = \alpha + \beta_0 y_{i,t} + \beta_1 y_{i,t-1} + \beta_2 y_{i,t-2} + \gamma Z_{i,t} + \xi Trend_{t,s} + \varepsilon_{i,t,s}, \quad (7)$$

and AIS equation,

$$y_{i,t,s} = \alpha + \beta_0 y_{i,t} + \beta_1 y_{i,t-1} + \beta_2 y_{i,t-2} + \gamma Z_{i,t} + \delta Q_{i,t,s} + \xi Trend_{t,s} + \varepsilon_{i,t,s}. \quad (8)$$

In the above equations, i indexes households, and $Z_{i,t}$ is a vector of variables that belongs to $\mathcal{I}_t^R$. We only include information about the household head.⁵ This vector contains race, education level, marital status, family size, a currently laid-off or unemployed dummy, a self-employed dummy, and a vector of occupation and industry dummies, all evaluated at age t . In addition $Z_{i,t}$ includes a fourth-order polynomial of imputed years of working experience, evaluated at the target age $t + s$.⁶ $Q_{i,t,s}$ is a vector of variables that belongs to the augmenting future information set, $\mathcal{I}_t^F$. We assume $Q_{i,t,s}$ includes family size, marital status, a retirement dummy, a part-time dummy, a self-employment dummy, and a vector of occupation and industry dummies, all evaluated at age $t + s$.

Besides $\mathcal{I}_t^R$ and $\mathcal{I}_t^A$, our specification departs from most previous specifications in that we include both current and lagged income in our projection equation. In principle, if we have a very long income history for a given household, even a univariate ARIMA model could potentially have decent forecasting power. Meanwhile, such a long time series of household income will be useful to identify the potential income process heterogeneity that Lillard and Weiss (1979) and Guvenen (2007) have stressed. Including some recent income history can help to tease out information about recent income shocks and capture part of the individual-specific information of income growth that is not revealed by current income. In practice, our model includes two lags to preserve degrees of freedom. Finally, we add a simple calendar year trend to control for aggregate economic growth.⁷

There are several important caveats regarding our model. First, because most households do have plans about their family and career ahead of

⁵Our data source does not have information for spouses that is as complete as heads. Including available spousal information does not significantly change the results.

⁶Experience is imputed as age minus 15, 18, 20, or 22 for the categories of high school drop out, high school graduate, some college education, and college graduate respectively.

⁷We also estimated the model with year dummies, assuming households have perfect foresight on future aggregate growth and business cycle fluctuations. The results are very much similar to the model with a linear trend.

time, it is not unreasonable to assume households know several years ahead of time what their family size and marital status will be; whether they will be working, retired, or self-employed; and whether they will change occupation and industry. However, it is more difficult to firmly argue that households know all this information when s is large.⁸ Second, we try our best to project $y_{t,s}$ on an information set as close to $\mathcal{I}_t^H$ as possible, subject to the data availability constraint. Nevertheless, it is still possible that there exist information elements ι such that $\iota \in \mathcal{I}_t^H$, but $\iota \notin \mathcal{I}_t^R \cup \mathcal{I}_t^F$ since econometricians only observe information collected by household surveys. However, as we will present later, the large value of R^2 of Eq.(8) reassures us that such omitted information can only have a limited impact. Third, Eqs. (7) and (8) are forecasting equations, and they are estimated solely on the basis of maximizing R^2 . We are not estimating a structural model, so the coefficients estimated for Eqs. (7) and (8) should not be interpreted as structural parameters. Finally, although we allow income uncertainty to vary with age, we compute these forecast errors by applying the same coefficients of the forecasting model to all households of different ages. As a robustness check, we estimate the forecasting model separately for each age group. Our results are qualitatively preserved, which reassures us that the age profile of income uncertainty is not driven by any age dependence in the accuracy of the forecasting model.

4 Data Description and Sample Construction

We use data from the Panel Study of Income Dynamics (PSID). The PSID is a nationwide longitudinal survey of households conducted by the Institute of Social Research at the University of Michigan. Before 1997, the PSID was an annual survey, while after 1997 it became a biannual survey. Currently, there are 34 waves of data that cover 38 years. The first wave of data was collected in 1968, and the latest available wave was collected in 2005. The PSID not only surveys the households in the original sample constructed in 1968, but also the households headed by the grown-up children of the original sample of households. The sample of households in the PSID has a very high retention rate. The vast majority of households that the PSID surveys in one year will continue to participate in the next wave. There are more than 1,200 households that stayed in the survey for more than 30 years. Consequently, the sample size of the survey has grown considerably since 1968. The first wave of the PSID had only 4,802 households, whereas the 1994 wave had more than 10,000 households. The PSID subsequently stopped surveying households in its non-core sample. As of the most recent wave of 2005, the survey has 8,002 households.

Besides extensive information about work status, employment history and demographic characteristics, the PSID has detailed information on house-

⁸Our study forecasts future income up to twenty-five years ahead.

hold income. In the current paper we report results regarding two income measures: total family income and household head labor earnings.⁹

The longitudinal structure of the PSID data allows us to link a household at time t to the same household at time $t+s$.¹⁰ However, we do not exploit other longitudinal properties of the data. For instance, consider a household that was surveyed over ten years from 1971 through 1980. When we project the five-year-ahead income, this household renders five current and future income pairs $(t, t+s) = (1971, 1976), \dots, (1975, 1980)$. We simply pool these pairs together without controlling for the household fixed effect.¹¹

Several selection rules apply when we construct the sample. First, we restrict the heads of our sample households to be those younger than 65 years old at the year to be forecasted. Thus, if we project income of year $t+s$ as of year t , we keep only the households whose heads are younger than $65-s$ as of year t . For example, in the sample we use to forecast five-year-ahead income, we restrict the heads of sample households to be younger than 60 years old. Consequently, the sample we use to study forecast errors at farther horizons is smaller than the sample used for a closer horizon. On the other side of the age restriction, we remove all households whose heads are younger than 23 years old in the base year. Second, we remove households that reported zero or negative income in the year to be forecasted. Third, we remove households whose heads are disabled or retired, are primarily keeping house, or are students in the base year. Fourth, we remove households whose heads report zero working hours in the base year and the previous two years or did not report valid industry or occupation information. Finally, in order to minimize the bias caused by outliers, we trim off households with very high or very low income levels and growth rates¹².

In our study, we estimate the variance of projection errors of incomes up to 25 years in the future. Because we include two lags of income in our forecasting equation and the PSID started in 1968, the first base year is 1970. We do not use the PSID 2003 and 2005 data because these waves used different occupation and industry codes that cannot be mapped to those used in the previous waves.¹³ Table 1 lists the number of observations used in the total family income estimation for each forecast horizon. The number of valid observations used in estimating the head labor earnings is somewhat smaller at all horizons. Table 2 provides summary statistics of key variables at the base year for the one-year-ahead sample, which is the largest. The income variables are deflated using 1982-1984 dollars.

⁹We also implement similar exercises with two more income measures, head wages and salaries and household labor earnings. The results are qualitatively similar.

¹⁰We abuse notation here to use t to denote age and time interchangeably.

¹¹As a robustness check, we also estimate the model controlling for clustering and the results are not much changed.

¹²We trim off the top and bottom 1% of lagged, current, and future income level distributions and the distribution of income growth between year t and $t+s$.

¹³All waves but 2003 and 2005 of the PSID data have 1970 census industry and occupation code. The 2003 and 2005 PSID data used the industry and occupation code derived from the 2000 census.

Table 1: Number of Observations for Each Sample

<i>s</i>	1 Year	2 Years	3 Years	4 Years	5Years
<i>N</i>	72,441	71,082	61,629	59,085	53,077
<i>s</i>	6 Years	7 Years	8 Years	9 Years	10 Years
<i>N</i>	49,004	44,037	39,851	35,782	31,935
<i>s</i>	11 Years	12 Years	13 Years	14 Years	15 Years
<i>N</i>	28,458	25,174	22,243	19,440	16,899
<i>s</i>	16 Years	17 Years	18 Years	19 Years	20 Years
<i>N</i>	14,515	12,495	10,547	8,991	7,419
<i>s</i>	21 Years	22 Years	23 Years	24 Years	25 Years
<i>N</i>	6,147	4,905	3,870	2,922	2,097

Table 2: Summary Statistics of the One-Year-Ahead Forecast Sample

Variable	Mean	StdDev	Variable	Mean	StdDev
Log(Family Income)	10.21	0.62	Log(Head Labor Earnings)	9.80	0.66
Headage	39.81	10.68	Family Size	3.31	1.75
< High School	29.1%		High School	23.3%	
Some College	29.4%		College	18.2%	
White	66.6%		Married	71.1%	

5 Main Empirical Results

We report five-year centered moving averages of the estimated forecast-error variances for households between 25 and 62 years old to further filter out possible noise in the series. Because we have from one- to twenty-five-year-ahead income projections, our income uncertainty matrix Ω has dimensions 38×25 .¹⁴ Figure 1 presents the nonparametric estimates of total family income uncertainties, whereas Figure 2 presents the counterpart results for household head labor earnings. For each income measure and model specification, we plot uncertainty profiles at four representative forecast horizons: one and two years (near future forecast), as well as five and ten years (intermediate and remote future forecast). To show the statistical significance of our results, for each profile, we also plot the 95-percent confidence interval band. In each figure, the top four panels are estimated using the AIS equation, and the bottom four panels are estimated using the RIS equation.

5.1 Total Family Income Uncertainty

We first focus on total family income uncertainties estimated using the AIS equation. These are displayed at representative horizons in the top four panels of Figure 1. For better contrast, we use the same vertical axis scale for the top two panels, which show uncertainty at one- and two-year horizons. Not surprisingly, the level of uncertainty regarding income two years ahead is on average 40% higher than the income uncertainty one year ahead. The confidence intervals associated with both uncertainty profiles indicate this gap is statistically significant. Apart from the level of uncertainty, both profiles share a similar U-shaped pattern over the lifecycle. Income uncertainty is at a high level in the mid twenties but declines during the late twenties and early thirties. Afterwards, uncertainty stays at a relatively stable level before rising again in the mid-forties. This rise continues as households approach retirement. The ratio between the maximum and minimum level of uncertainties over the lifecycle (the max-min ratio) is 1.46 for the one-year-ahead forecast and 1.35 for the two-year ahead forecast. We note that both profiles are tightly estimated, which reinforces the significance of the U-shaped pattern. A similar convex profile of income uncertainty is also documented in Baker and Solon (2003). They decompose the stochastic part of income into a transitory and a permanent component and find that the age profile of the variance of the transitory innovation exhibits a pronounced U-shaped pattern.¹⁵

Profiles of similar shapes repeat in the two panels on the second row, which show income uncertainties at the intermediate and remote future. The standard errors are somewhat larger than in the top two panels because we have a smaller sample over these longer horizons. The uncertainty associated with income five years ahead hits bottom in the mid thirties and stays quite flat until

¹⁴The lower-right triangle of the matrix is populated with zeros.

¹⁵A similar pattern was also noticed by Gordon (1984).

the mid forties. The uncertainty profile of income ten years ahead also exhibits a like pattern, but it bottoms at a somewhat earlier age. Adjusted for different forecast horizons, the bottom of each curve corresponds to a target age in the early forties. Comparing with the near future, the dynamics of the intermediate and remote future income uncertainty are less pronounced. The max-min ratio is 1.27 for the five-year horizon and 1.20 for the ten-year horizon. Both ratios are appreciably lower than the max-min ratios at nearer horizons. Moreover, though the level of uncertainty in the intermediate future is much higher than in the near future, the increase is not linear in forecast horizons. The two-year ahead uncertainty is 40% higher than the one-year ahead uncertainty. If these uncertainties exhibited a linear trend, the five-year ahead uncertainty, for example, should be 160% higher than the one-year ahead uncertainty. However, it is only 100% larger.

Finally, income-uncertainty profiles at the very remote future such as beyond twenty years were also estimated. Because of smaller sample sizes at these horizons, these profiles are less precisely estimated, so we do not report them in the figures. The point estimates also exhibit a U-shaped pattern, but these are not statistically significant.

Now we turn our attention to family income uncertainties as estimated under the RIS, which are shown in the bottom four panels of Figure 1. Several features of this figure are noteworthy. First, the U-shaped contours of these lifecycle uncertainty profiles are qualitatively similar to those in the AIS results, shown in the top four panels. Second, the max-min ratios are somewhat higher for the RIS estimates, especially for income in the near future. Third, because these future income projections are conditioned on a more restricted information set, the RIS forecast error variances are greater than the variances estimated under AIS. This is true for all forecast horizons and for households of all ages. Finally, as we expect, the discrepancy between the RIS and AIS results widens with the forecast horizon. The RIS one-year-ahead uncertainty is only 4% higher than the AIS uncertainty, whereas the margin is about 22% at the ten-year horizon. This is because the difference between $\mathcal{I}^R$ and $\mathcal{I}^A$ is what households might know in the base year but what econometricians only observe after a s -year lag. If s is small, the correlation between the elements in $\mathcal{I}^R$ and $\mathcal{I}^F$ will be high, discounting the net value of augmenting with $\mathcal{I}^F$. Conversely, if s is large, $\mathcal{I}^R$ should have less predictive power on $\mathcal{I}^F$. Consequently, augmenting by $\mathcal{I}^F$ adds much more new information and beefs up the forecasting performance.

5.2 Household Head Labor Earnings Uncertainty

Figure 2 shows the corresponding results for the uncertainty of household head labor earnings. Because we do not include any spousal information in the projection equation, using head income is conceptually more consistent with the model specification, although numerically the contribution of spousal information is small. Comparing to the total family income uncertainty profiles, the head labor earnings profiles share many common features. First, they evolve over the lifecycle with similar dynamics. Second, uncertainty levels increase

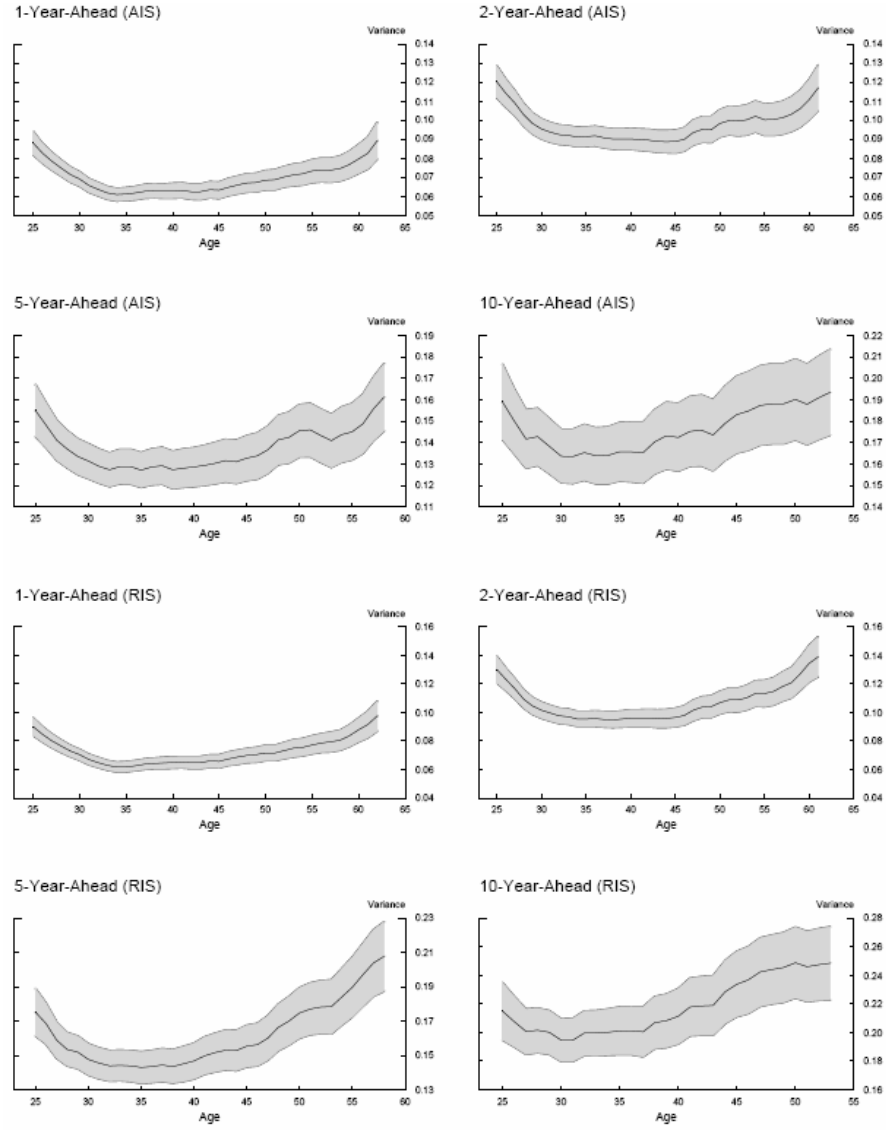


Figure 1: Lifecycle income uncertainty profiles of total family income. The top four panels are estimated using the AIS equation, and the bottom four panels are estimated using the RIS equation. The shaded areas represent 95-percent confidence intervals.

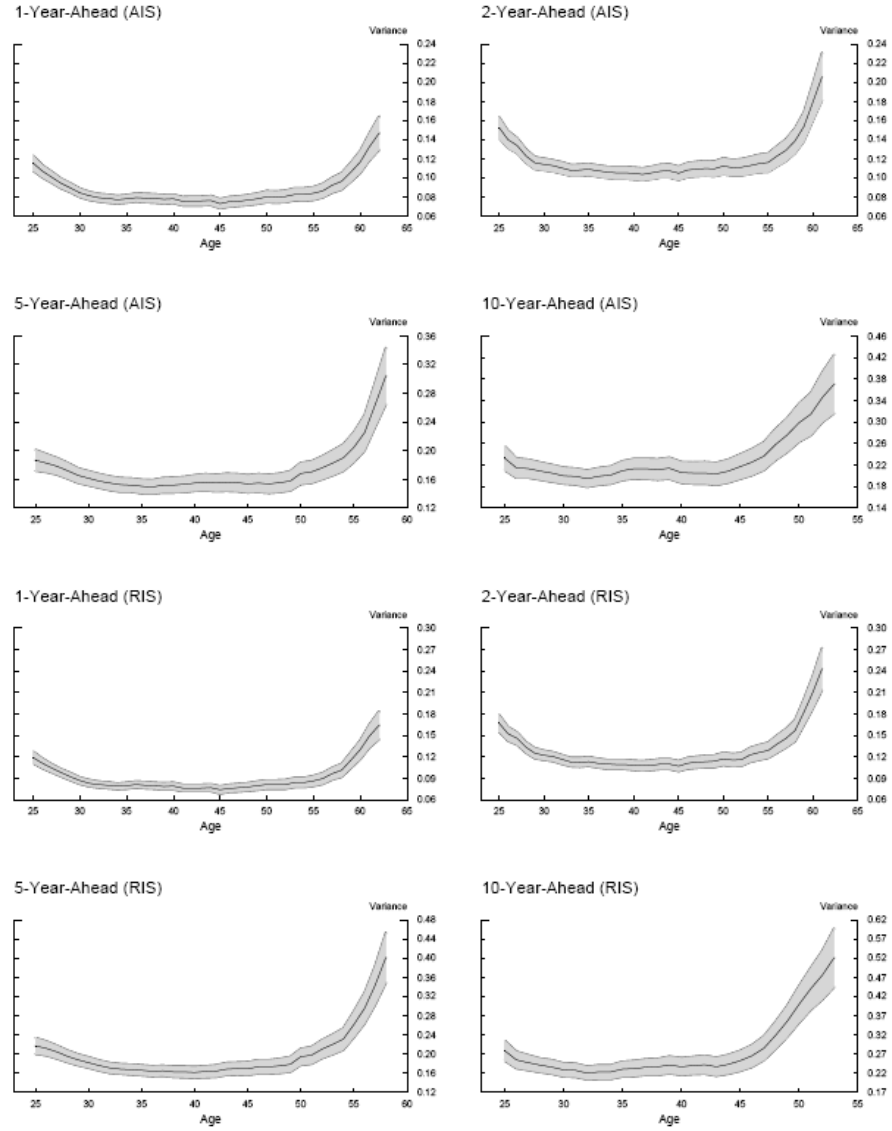


Figure 2: Lifecycle income uncertainty profile of labor earnings of household heads. The top four panels are estimated using the AIS equation, and the bottom four panels are estimated using the RIS equation. The shaded areas represent 95-percent confidence intervals.

with the forecast horizon. Third, the gap between the AIS and RIS estimates also ranges from 4% for income in the near future to about 22% for income in the remote future.

The most important distinction between the uncertainty dynamics of the two income measures is that the rise of uncertainty in the later phase of the lifecycle is much more pronounced for head labor earnings than for total family income. The former evolves with a U-shaped pattern, whereas the latter evolves with more of a J-shaped pattern. Taking the ten-year-ahead uncertainties as an example, this series covers households between 25 to 53 years old. Under the AIS, the uncertainty for total family income ten years in the future perceived at age 25 is about the same as what is perceived at age 53. In contrast, for head labor earnings, the uncertainty perceived at age 53 is almost 50% higher than what is perceived at age 23, suggesting that either households face substantially greater head labor earnings risk than total family income risk later in life, or our forecast equation does a better job capturing predictable variation for total family income than for head labor earnings.

5.3 Correlations of Forecast Errors

Besides the magnitude of income uncertainties over the lifecycle, to completely characterize the dynamics of income uncertainty we also must know how the stochastic components of income at different horizons correlate. We present for each age group the correlations between the forecast errors one-year ahead and the forecast errors at other horizons (Θ^1) in Figure 3. To keep the graph readable, we only show the correlations of total family income uncertainties faced by households that are thirty, forty, and fifty years old under the AIS. Correlations for other income measures and for households of other ages are similar. The chart shows that, apart from the longest forecast horizons when sample sizes are smallest, the correlations of households of various ages follow the same pattern. Since the correlations are positive at all forecast horizons in our sample, income shocks must have a persistent component. We note that the correlations decline with the forecast horizon. The correlation between the one-year and two-year forecast errors is about 0.45 whereas the correlation between the one-year and ten-year forecast errors is only about 0.15, one third of the previous value. This rapid falloff suggests that the persistent component of income shocks is more likely to be an AR process than a unit-root process. We cannot rule out the existence of a unit root, though, if there is one, its innovations should have a relatively small variance.

5.4 The (In)significance of Lagged Income

One innovation of our model is that future income is projected conditional on not only current income, but also lagged income. The motivation for including lagged income in the projection equation is to capture some income-profile heterogeneity that cannot be accounted for by a common trend determined by

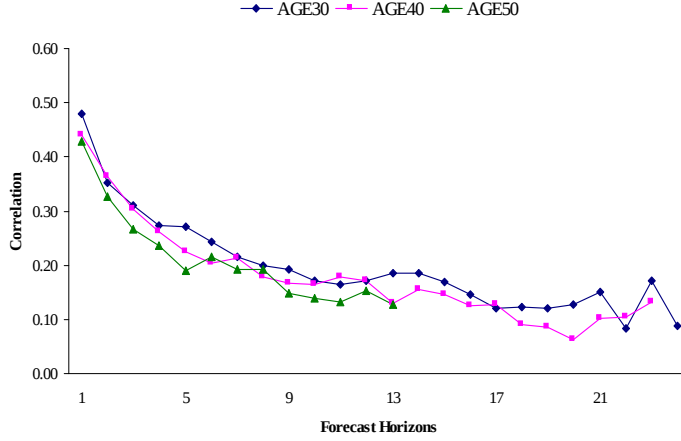


Figure 3: Correlations of 1-year-ahead and h -year-ahead forecast errors for ages 30, 40 and 50.

demographic, industry and occupational variables. This method will not capture all the heterogeneity set forth by Lillard and Weiss, but does acknowledge that past income history is useful in predicting future income. Indeed, for almost all income measures and forecasting horizons, the estimated coefficients of the lagged income are highly significant (with p -value typically smaller than 0.001). In addition, if we add up the current and lagged income coefficients, the sum is larger than the coefficient of the current income estimated in a model with no lagged income included, confirming that with lagged income included, the projection picks up more signal from the income history.

However, what is somewhat puzzling to us is that although the coefficients of the two lagged incomes are statistically significant, adding them does not beef up the overall fitness of the projection equation to a great extent. For example, the R^2 of the AIS equation for the total family income one year ahead merely increases from 0.7489 to 0.7612. Furthermore, the improvement of R^2 diminishes as forecast horizons become remote. This limited boost of fitness manifests in the variance of forecast residuals as well. As Figure 4 shows, including two lags of income reduces the one-year-ahead forecast residual variance by about 5.5 percent in the AIS model. This gap decreases to only 1.2 percent

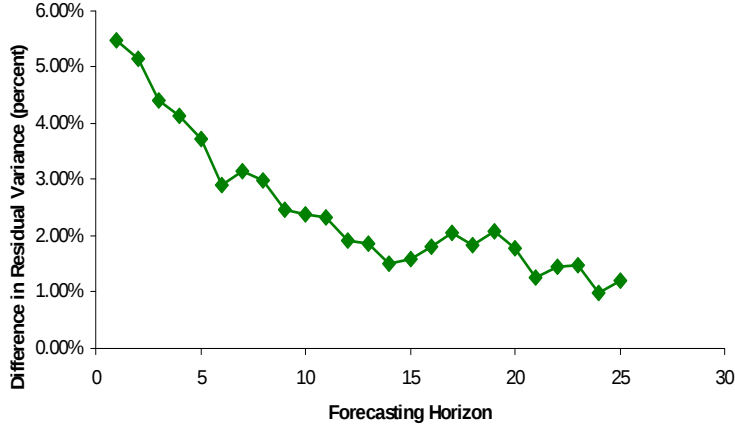


Figure 4: Contribution of including lagged income in the projection equations.

at the twenty-five-year horizon.

On balance, we find that including a short income history can moderately boost forecast performance in the near future and does not have a significant effect in the remote future. However, this result does not imply that income heterogeneity bears no significance. Rather, the insignificance of lagged income suggests that much of the heterogeneity can only be accounted for by adding more information beyond a short income history.

5.5 Robustness Tests

To verify whether the U-shaped uncertainty profile over the lifecycle is a spurious consequence of our model specification, sample size, or sample selection, we conduct a series of robustness tests. First, we examine whether the changes in forecast-error variances are due to model misspecification. Remember we project the future income of households of different head ages using the same set of coefficients. If the projection-equation coefficients should be age-specific and the projection equation we use is closer to the true equations for middle-aged households than to the true equations for younger and older households, the one-size-fit-all approach will reduce fitness for younger and older households

and artificially increase income uncertainties for these age groups. We divide our sample into five subgroups by head of household age and reestimate Eqs. (7) and (8) separately for each subgroup. Then we calculate forecast-error variances as we did before and we find the U-shaped uncertainty profiles are preserved, though with wider confidence intervals.

Second, we examine whether changes in the sample size as we vary the forecast horizon (as given in Table 1) might drive the shape of the uncertainty profile. We reestimate the forecast equations using a smaller common sample and reassess income uncertainties over the lifecycle. Apart from the fact that more matrix elements cannot be estimated accurately because of the smaller sample size, the magnitude and dynamics of income uncertainty are similar to what we presented above.

Finally, we test if our results are driven by low-income households, which are oversampled by the PSID. The core PSID sample consisted of two independent samples: a nationwide representative sample and a sample of low-income families. In the first wave of the survey, the nationwide representative sample has about 3,000 households and the low-income sample has about 2,000 households. We redo our analysis using only the households in the nationwide representative sample and their offspring. The results are similar to those obtained using the entire PSID core sample.¹⁶

6 Discussion and Comparison with Earlier Results

How substantive are the innovations we have introduced into these nonparametric measures of income uncertainty? How different are our results compared to previous results in the literature? We answer these questions by contrasting our results to the income uncertainty estimates in the influential work of Carroll and Samwick (1997) and Gourinchas and Parker (2002). Three reasons lead us to choose their results to compare with. First, their specific interest was precautionary saving, so their estimates focus explicitly on income uncertainty and risk rather than inequality, volatility, or other types of income variability. Second, their work also use the PSID data, which makes it easier to update their results to implement a comparison and to interpret the outcome. Third, apart from the specified innovations, our model specifications share many similarities with theirs.

As we summarized in Eqs. (1)-(2) with ρ set to 1, Carroll and Samwick (1997) and Gourinchas and Parker (2002) decomposed the logarithm of income, y_t , into a permanent component, p_t , and a transitory shock, ϵ_t . The permanent component, p_t is further assumed to follow a random walk with predictable

¹⁶For more information about the PSID sample design, see the online documentation at <http://psidonline.isr.umich.edu/Guide/Overview.html>.

income growth g_t such that

$$p_t = g_t + p_{t-1} + \eta_t, \quad (9)$$

where η_t is the shock to permanent income. Let σ_η^2 and σ_ϵ^2 be the variance of the permanent and transitory shocks. Define VAR_d as the variance of the d -year income difference. Filtering out g_t , it is easy to show that

$$VAR_d = \text{Var}[y_{t+d} - y_t] = d\sigma_\eta^2 + 2\sigma_\epsilon^2, \quad (10)$$

noting that the econometrician does not know how either y_t or y_{t+d} decomposes into their permanent and transitory parts. The innovation variances σ_η^2 and σ_ϵ^2 can be estimated by evaluating $\text{Var}[y_{t+d} - y_t]$ at various difference lengths, d . Using a PSID sample from 1981 to 1987, Carroll and Samwick report $\sigma_\eta^2 = 0.022$, and $\sigma_\epsilon^2 = 0.044$. Gourinchas and Parker (2002) report almost identical results. These parameters are changed significantly when we estimate their model using an extended PSID sample that covers a longer period 1968 - 2001. The updated estimates call for a much bigger variance for the transitory income shocks, $\sigma_\epsilon^2 = 0.071$, and a much smaller variance of permanent shocks, $\sigma_\eta^2 = 0.009$.

Figure 5 contrasts the total family income uncertainty at various forecast horizons implied by the original and updated Carroll and Samwick estimates with our nonparametric profiles (pooled across ages). The higher and steeper linear profile is implied by the σ_η^2 and σ_ϵ^2 reported in Carroll and Samwick (1997), the lower and flatter linear profile is implied by the σ_η^2 and σ_ϵ^2 estimated using the extended PSID sample. The higher concave profile is estimated using the RIS equation, whereas the lower concave profile is estimated using the AIS equation. The linearity between income uncertainties and forecast horizons under the Carroll and Samwick specification arises because of the random walk assumption imposed on the permanent income shocks. The slope of the linear profile is equal to the variance of permanent shocks, and the intercept is equal to two times the variance of transitory shocks.

There are several noteworthy features in this chart. First, we notice that the Carroll-Samwick (1997) profile is substantially higher than both nonparametric profiles estimated under the AIS and RIS equations. At near horizons, the gap is between 20-40%, whereas at remote horizons, because both AIS and RIS nonparametric profiles are concave, the wedge between the linear profile and the nonparametric estimate widens substantially to between 70-120%. Second, the updated Carroll-Samwick profile has a higher intercept but a flatter slope due to the larger σ_ϵ^2 but smaller σ_η^2 . Third, the predictable income component, $g_{t,s}$, is constructed using time $t + s$ information in the Carroll and Samwick approach, whereas the RIS equation is conditioned only on time t information.¹⁷ Thus it is very striking that even the updated Carroll-Samwick estimates can not outperform the RIS estimates.¹⁸

Why are the uncertainty profiles implied by the original and updated

¹⁷For the specific application that Carroll and Samwick (1997) studied, they sought to measure income uncertainty across the population at a specific point in time, corresponding

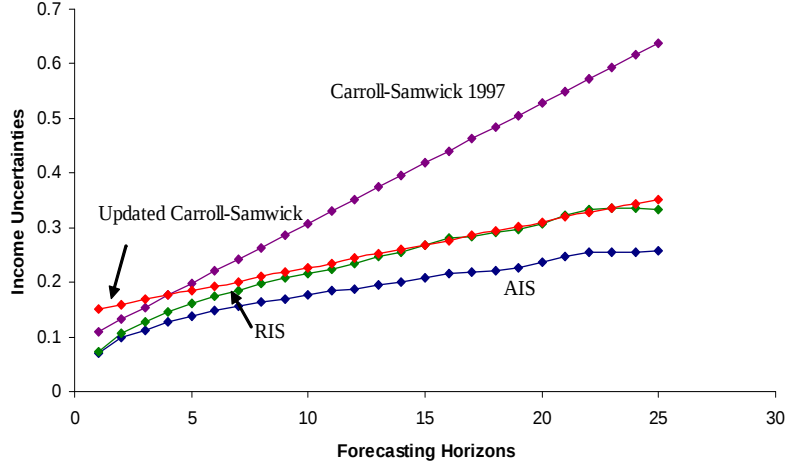


Figure 5: Camprison between the implied total family income uncertainty at various future horizons implied by four estimates: (1) the parameters reported in Carroll and Samwick (1997); (2) the updated Carroll and Samwick estimates using an extended PSID sample; (3) the AIS estimates; and (4) the RIS estimates.

Carroll and Samwick estimates so different? Two potential factors may contribute to this change. We notice that in Carroll and Samwick (1997), the maximum difference length corresponds to $d = 6$, which is relatively small. Using a longer panel not only adds data from more years, but it also allows us to consider the difference in income over longer intervals. To assure the comparability of our PSID sample with theirs, we first replicate their results using the PSID 1981-1987 sample that we construct. Our estimated transitory shock variance is 0.044, identical to theirs, and our estimated permanent shock variance is 0.019. The latter variance is slightly smaller than theirs, but the gap is not statistically significant. Then we update these estimates using the extended PSID sample. We vary $\max[d]$ to examine whether the estimates are sensitive

to $t + s$ in our notation.

¹⁸The updated Carroll-Samwick profile is also uniformly higher than the AIS profile at all forecast horizons by a margin of 30% on average. However, we acknowledge that the AIS estimates condition on both time t and $t + s$ information.

to variation in difference length.

Figure 6 contrasts how the estimated variances of transitory and permanent shocks, identified by the specification of Eqs. (9)-(10), vary with $\max[d]$. Because we project future income up to 25 years ahead in our nonparametric estimates, we choose the largest value of $\max[d]$ to be 25. This figure has several outstanding features. First, the permanent shock variance decreases monotonically with $\max[d]$ whereas the transitory shock variance increases. This phenomenon implies either that the variance of the permanent shocks—if the persistent shock process is indeed a random walk—is estimated with a significant upward bias using a short panel, or that the persistent shocks do not follow a random walk but rather an AR process. To see this, suppose the true model of the shocks to “permanent” income is

$$p_t = \rho p_{t-1} + \eta_t, \quad (11)$$

instead of a unit root process $p_t = p_{t-1} + \eta_t$. We can show that, after some algebra, VAR_d is not equal to $d\sigma_\eta^2 + 2\sigma_\varepsilon^2$ as in Eq. (10). Rather, we have

$$VAR_d = \text{Var}[y_{t+d} - y_t | y_t] = \frac{1 - \rho^{2d}}{1 - \rho^2} \sigma_\eta^2 + [1 + \rho^{2d}] \sigma_\varepsilon^2. \quad (12)$$

Using L'Hôpital's Rule, it is easy to verify that (10) is the limiting case when $\rho \rightarrow 1$. If eq. (10) holds, σ_η^2 can be calculated by taking the difference $VAR_{d+1} - VAR_d$. The efficient estimate that Carroll and Samwick implemented can be viewed as a weighted average of such variances across $d \leq \max[d]$. However, if Eq. (12) is the true model, taking the difference between VAR_{d+1} and VAR_d , we get the presumed estimate of the variance of permanent shocks,

$$VAR_{d+1} - VAR_d = \rho^{2d} [\sigma_\eta^2 + (\rho^2 - 1) \sigma_\varepsilon^2]. \quad (13)$$

Assuming ρ is close to 1 (e.g. $\rho = 0.9$) and σ_η^2 and σ_ε^2 are of the same order of magnitude, then $VAR_{d+1} - VAR_d$ will be a decreasing function of d .¹⁹ Therefore, when we increase $\max[d]$, the average of $VAR_{d+1} - VAR_d$ over $d \leq \max[d]$ also decreases. Meanwhile, a downward biased estimate of the permanent shock variance leads to an upward biased estimate of the transitory shock variance.

Second, when we restrict $\max[d] = 6$ in the extended PSID sample and throw out longer differences, the estimated permanent shock variance is 0.015 and the estimated transitory shock variance is 0.054. The former is smaller than the 1981-1987 PSID sample estimate, whereas the latter is larger. This could be because the variance of transitory shocks increased in later samples. This explanation is consistent with the notion that the contributions of transitory shocks to changes in income inequality have become increasingly more significant over the last three decades, as argued by Gottschalk and Moffitt (1994). However, the current paper is focusing on the dynamics of income uncertainty

¹⁹In Section 7, we calibrate $\sigma_\eta^2 \approx \sigma_\varepsilon^2 \approx 0.04$ for an analogous experiment with an AR persistent shock.

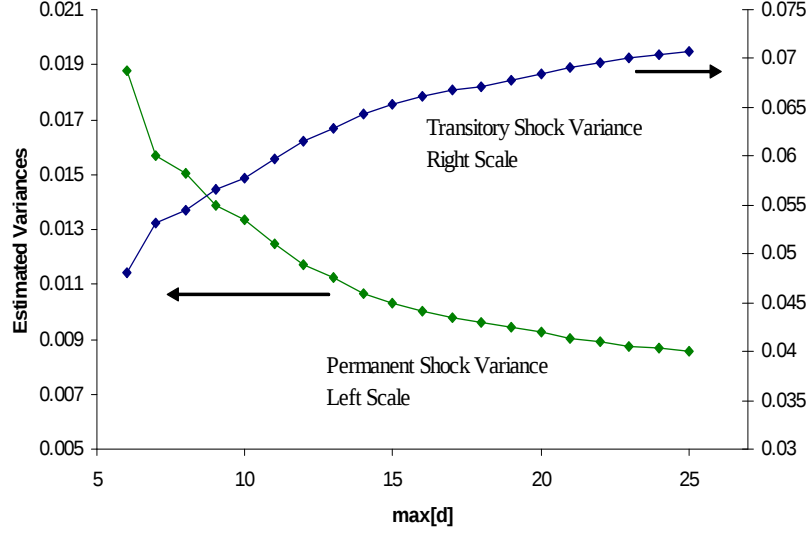


Figure 6: Variations in the estimated variances of transitory and permanent income shocks as a function of the maximum time lag.

over the lifecycle. We leave formal study of the dynamics over calendar time to future work.

Finally, Carroll and Samwick (1997) also estimate the variance of the permanent and transitory income shocks by age. However, their estimates do not imply a U-shaped lifecycle pattern. Indeed, their transitory shock variance exhibits a hump-shaped pattern over the lifecycle, peaking in the early forties, whereas their permanent shock variance demonstrates more irregular lifecycle dynamics.

To summarize, in contrast to Carroll and Samwick (1997) and later work using similar methods, our nonparametric estimates of income uncertainty reveal a lower level of income uncertainty and suggest that the persistent component of income shocks is not a permanent shock. Our approach also implies lifecycle dynamics of income uncertainty more consistent with layman's intuition and the results of Baker and Solon (2003) and Gordon (1984).

7 Calibration

While our nonparametric estimates of the volatility and correlation matrices fully characterize the second moments of income forecast errors, for many theoretical applications it is necessary to specify a complete data generating process for income. For example, to study the impact of uncertainty on consumption and saving, we need the data generating process to compute optimal consumption functions over the course of the lifecycle. As Guvenen (2007) has recently pointed out, there is considerable disagreement about the proper specification of individual income processes, so rather than estimate a parametric income process we choose to calibrate a flexible yet computationally tractable process that is common in the literature. Thus we assume that income decomposes into a deterministic age-dependent component, a persistent shock, and a temporary shock. Carroll and Samwick (1997) and Gourinchas and Parker (2002) also employ such a decomposition, but we will deviate from their approach in two important respects. First, given that our nonparametric estimates call for a concave increase of income uncertainty with forecast horizon, as opposed to a linear trend, the persistent shock in our specification will not follow a unit-root process.²⁰ Second, the variance of the persistent and temporary shocks will depend on age and, in the case of the temporary shocks, forecasting horizon.

Formally, we decompose log income at age $t + s$, conditional on information available at age t , as

$$y_{t,s} = a_{t+s} + p_{t,s} + z_{t,s}, \quad (14)$$

where a_{t+s} is the age-dependent factor, $p_{t,s}$ is the persistent shock, and $z_{t,s}$ is the temporary shock. More precisely, we assume that $p_{t,s}$ and $z_{t,s}$ are independent, unconditionally unit-mean processes such that

$$\begin{aligned} p_{t,s+1} &= \rho p_{t,s} + \sigma_p(t) \eta_{t,s} \\ z_{t,s} &= \sigma_{z,s}(t) \varepsilon_{t,s}, \end{aligned}$$

where $\eta_{t,s}$ and $\varepsilon_{t,s}$ are i.i.d. zero-mean and unit-variance white noise processes, which are also independent of each other. Note that we assume the age-dependent factor a_{t+s} and the autocorrelation ρ are the same for each base age t .

For comparison, we also consider a model more typical of the existing literature, where the variance of the persistent and temporary shocks is simply a constant, independent of time. The age-dependent and constant-variance calibrations of $\sigma_p(t)$ are plotted as a function of age t in Fig. 7 for both the AIS and RIS specifications.²¹ Likewise, the calibrations of $\sigma_{z,h}(t)$ are plotted

²⁰Note that many papers in the literature have previously considered an AR(1) process for the persistent shocks, including Feigenbaum (2008) and Huggett (1996).

²¹For details of the calibration procedure, see Feigenbaum and Li (2008), which refers to the age-dependent income process as the time-inconsistent model and the constant-variance process as the time-consistent model.

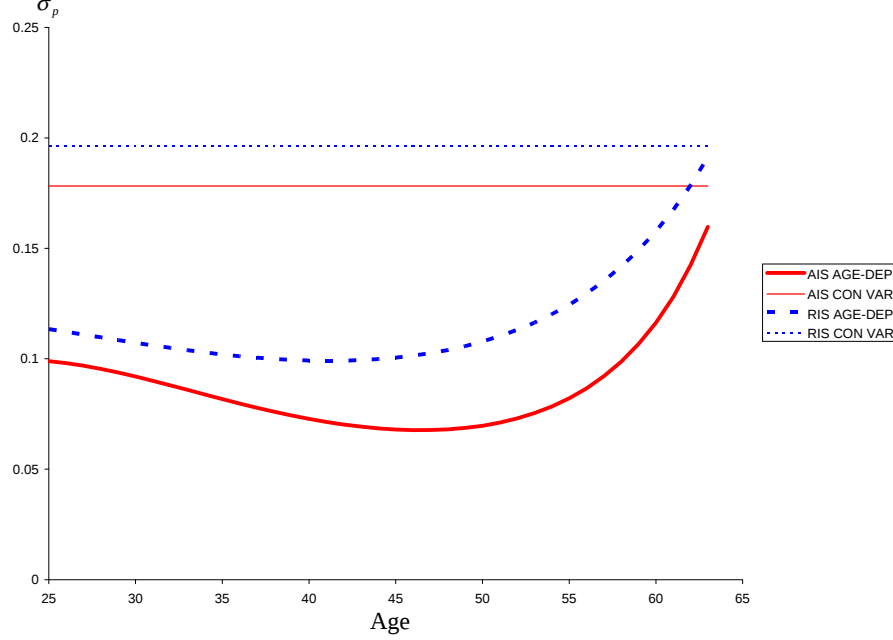


Figure 7: Persistent shock standard deviation $\sigma_p(t)$ as a function of age t for the age-dependent and constant-variance income processes with both the RIS and AIS estimates of the volatility and correlation matrices.

as a function of age for four representative horizons for each of the four models in Fig. 8. The correlation coefficients for each calibration are given in Table 3.

The variance of the persistent shock is larger at all ages for the constant-variance calibration than the age-dependent calibration. At short time horizons the variance of the temporary shock is comparable between the age-dependent and constant-variance calibrations. However, the variance of the temporary shock increases with the forecast horizon in the age-dependent model while necessarily remaining constant in the time-consistent model. As a result, for most horizons the variance of the temporary shock is twice as large for the age-dependent calibration as for the constant-variance calibration. Thus the persistent income shocks will have greater emphasis in the constant-variance calibration whereas temporary income shocks will have more emphasis in the age-dependent calibration. Autocorrelations for the constant-variance calibrations are modestly smaller than the corresponding age-dependent calibrations. This difference in autocorrelations may help account for the difference in the persistent-shock variances.

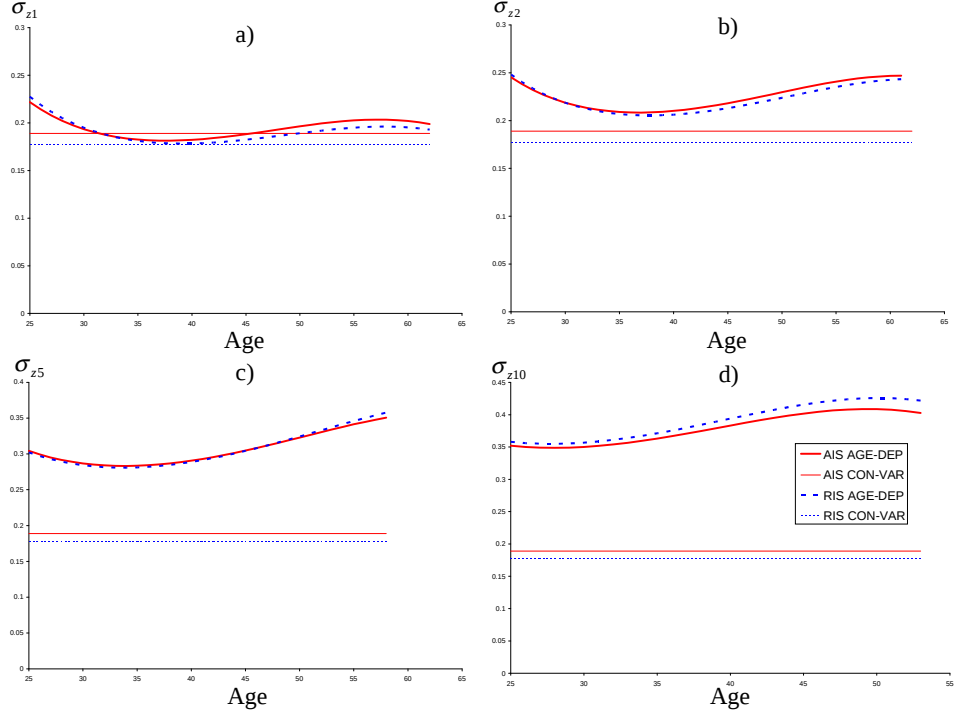


Figure 8: Temporary shock standard deviation $\sigma_{zh}(t)$ as a function of age t for horizons h of (a) one year, (b) two years, (c) five years, and (d) ten years for both the age-dependent and constant-variance income processes and both the AIS and RIS estimates of the volatility and correlation matrices.

Specification	ρ
AIS CON VAR	0.910
AIS TIME-DEP	0.958
RIS CON VAR	0.925
RIS TIME-DEP	0.964

Table 3: Correlation ρ for both the age-depeneent and constant-variance income processes and both the AIS and RIS estimates of the volatility and correlation matrices.

To see the quantitative significance of making the shock variances age-dependent, let us compare the root-mean-squared deviation between the predictions of each calibration and the corresponding nonparametric estimates since this is the metric we minimize in our calibration exercises. For the AIS estimates the root-mean-squared deviation is 0.022 whereas for the constant-variance calibration it is 0.051.²² Figs. 9-10 shows how the variances of forecast errors for these two models compare to the nonparametric estimates as a function of age at different forecast horizons. Since the constant-variance model assumes that the standard deviation of temporary and permanent income shocks is independent of age, its variance graphs are flat whereas the age-dependent model is able to capture the U-shaped dependence of the variances with respect to age. For the AIS, at short horizons of one to two years, the constant-variance model overpredicts the uncertainty at all ages. For intermediate horizons, the constant-variance model overpredicts the uncertainty in the 20s and 30s.²³ Consistent with its smaller root-mean-squared deviation, the age-dependent model almost uniformly does better at matching the forecast-error variances, although it does underpredict the variance at the two-year horizon. The comparison is similar for the RIS specification.²⁴

To get a better sense of the relative importance of the age-dependence of the income shock variances and forecast-horizon dependence, Fig. 11 compares the one-year ahead AIS estimates of uncertainty to the uncertainties obtained from the four models that arise depending on whether we allow the shock variances to depend on age or not and whether we allow the variances to depend on forecast horizon or not.²⁵ The figure shows that both age and forecast-horizon dependence are needed to fit the observed income uncertainties and, in particular, the U-shaped lifecycle profile of uncertainty. However, the dependence on forecast horizon is more crucial. The uncertainties that we obtain if we allow the shock variances to depend on age but not forecast horizon overpredict the observed uncertainties about as badly as the constant-variance model, and the root-mean squared error of 0.050 for the age-only-dependent model is roughly the same as for the constant-variance model. In contrast, while the uncertainties of the forecast-horizon-only-dependent model are flat by construction as a function of age, this model predicts one-year-ahead uncertainties comparable to the lifecycle average for these uncertainties, and the root-mean squared error of 0.028 is much closer to that of the age and horizon-dependent model than the constant-variance model. The dependence of the temporary shock variances on forecast horizon indicate that consumers do learn a great deal about their future income as they get closer to this future, but much of this information is not coming from their income history.

²²Note that if we imposed $\rho = 1$ in the constant-variance calibration, we would get a root-mean-squared deviation of 0.115 for the AIS estimates.

²³The constant-variance model significantly underpredicts the point estimates of uncertainty at long horizons of 15-25 years.

²⁴We do not report results for longer horizons because the variances are less precisely estimated, but we still obtain a good fit between the estimated lifecycle uncertainty profile and the one implied by the age-dependent calibration at these horizons.

²⁵We get similar results for other horizons.

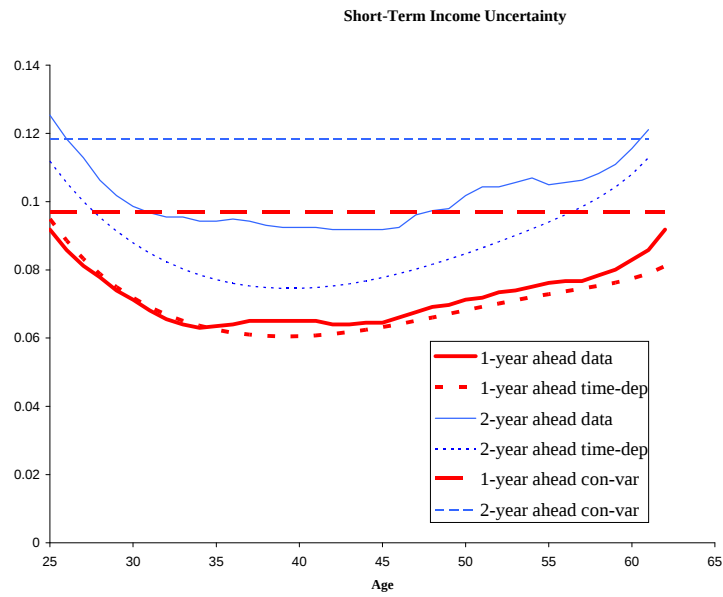


Figure 9: Variance of forecast errors at short horizons as a function of age for the augmented information set (AIS) estimates and both the constant-variance and age-dependent income processes.

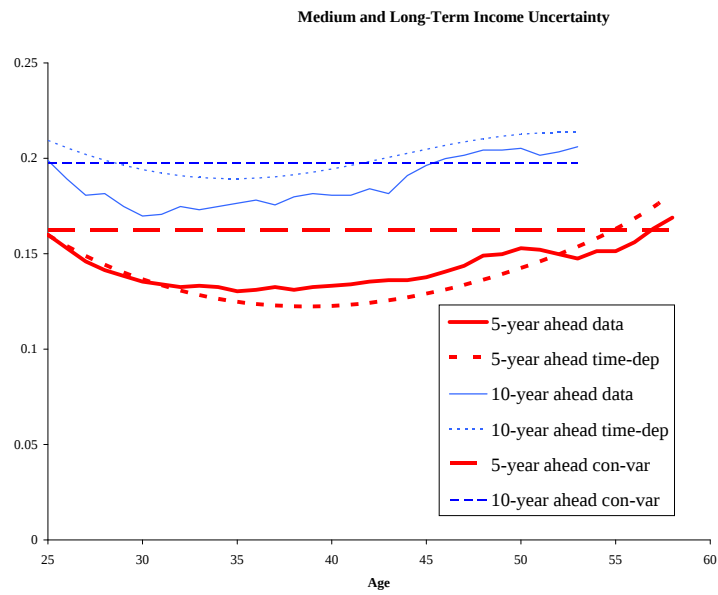


Figure 10: Variance of forecast errors at intermediate and long horizons as a function of age for the augmented information set (AIS) estimates and both the constant-variance and age-dependent income processes.

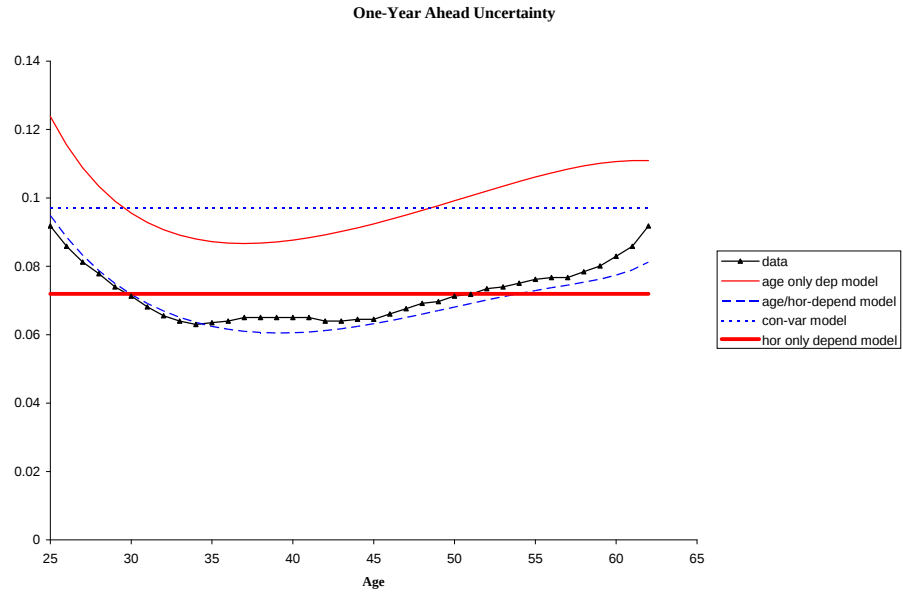


Figure 11: Variance of one-year ahead forecast errors as a function of age for the augmented information set (AIS) estimates and for the four possible models that arise based on whether variances are allowed to depend on age and whether they are allowed to depend on forecast horizon. Note that “age/hor-dependent model” corresponds to what we normally refer to as the age-dependent model in the text.

Finally, Fig. 12 shows the correlations between one-year ahead forecast errors and h -year ahead forecast errors. We plot the correlation for households at ages 30, 40, and 50 as estimated nonparametrically under the AIS specification and as implied by the age-dependent calibration. We also plot the correlation for the constant-variance calibration, which by construction is the same for all base ages. At all three of these ages, the age-dependent model matches the correlations better since the constant-variance model overpredicts the correlation at short horizons and underpredicts it at long horizons.

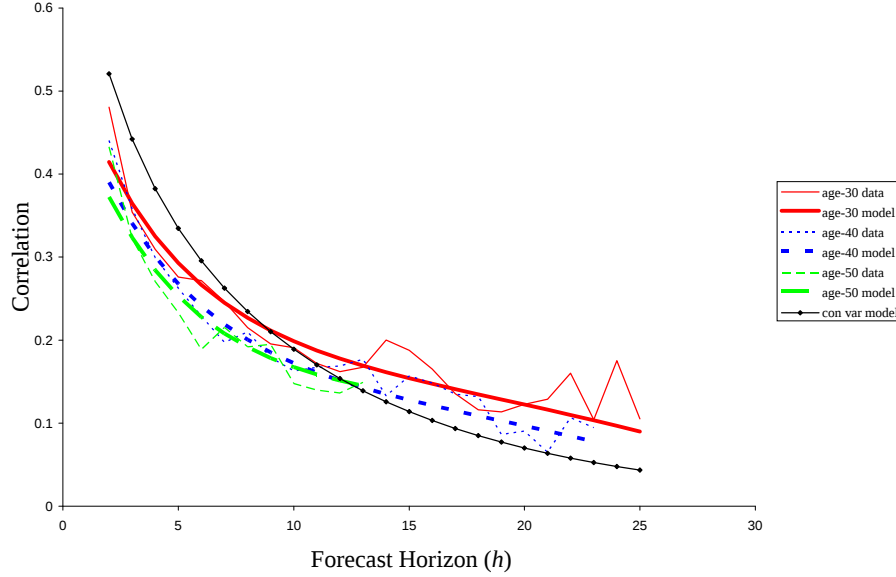


Figure 12: Correlation of one-year ahead forecast errors and h -year ahead forecast errors as a function of forecast horizon h at ages 30, 40, and 50 for the augmented information set (AIS) estimates and both the age-dependent and constant-variance income processes.

8 Conclusion

We construct a nonparametric measure of income uncertainty and study its dynamics over the lifecycle. Our estimates of income uncertainty are typi-

cally smaller than previous studies have documented. Our estimates also imply less persistence in income shocks. Over the lifecycle, we find robust U-shaped patterns in the evolution of income uncertainty. Young and old consumers on average have more risky future income relative to middle-age consumers. This U-shaped pattern is robust to a number of sample and model specifications and prevails at almost all horizons.

With this refined measurement of income uncertainty over the lifecycle, a wide variety of theoretical questions can be revisited, covering topics such as consumption and finance. For instance, Feigenbaum and Li (2008), using the income process that we calibrate here, show that precautionary saving can create a hump-shaped lifecycle profile of mean consumption that matches what is observed in the data without resorting to controversial preference parameters.

We can also employ the methodology introduced here to consider the evolution of income uncertainty for households of a fixed age as time progresses. Doing so will let us address the question of whether the Great Moderation people find in aggregate income data (Kim and Nelson (1999), McConnell and Perez-Quiros (2000)) manifests itself at the household level.

References

- [1] Aiyagari, S. Rao, (1994), "Uninsured Idiosyncratic Risk and Aggregate Saving," *Quarterly Journal of Economics* 109: 659-684.
- [2] Baker, Michael and Gary Solon, (2003), "Earnings Dynamics and Inequality among Canadian Men, 1976-1992: Evidence from Longitudinal Income Tax Records," *Journal of Labor Economics* 21: 289-321.
- [3] Campbell John Y. and Joao F. Cocco. (2003). "Household Risk Management And Optimal Mortgage Choice," *The Quarterly Journal of Economics* 118, 1449-1494.
- [4] Carroll, Christopher D., (1997), "Buffer-Stock Saving and the Life-Cycle/Permanent Income Hypothesis," *Quarterly Journal of Economics* 112: 1-55.
- [5] Carroll, Christopher D. and Andrew A. Samwick, (1997), "The Nature of Precautionary Wealth," *Journal of Monetary Economics* 40: 41-71.
- [6] Carroll, Christopher D. and Andrew A. Samwick, (1998), "How Important is Precautionary Saving?" *Review of Economics and Statistics* 80(3): 410-419.
- [7] Cunha, Flavio, James Heckman, and Salvador Navarro, (2005), "Separating Uncertainty from Heterogeneity in Life Cycle Earnings," *Oxford Economic Papers* 57: 191-261.

- [8] Dynan, Karen, Douglas W. Elmendorf, and Daniel E. Sichel, (2006), "Can Financial Innovation Help to Explain the Reduced Volatility of Economic Activity?" *Journal of Monetary Economics* 53: 123-150.
- [9] Dynan, Karen, Douglas W. Elmendorf and Daniel E. Sichel, (2007), "The Evolution of Household Income Volatility," Federal Reserve Board Finance and Economics Discussion Series 2007-61.
- [10] Feigenbaum, James, (2007), "Precautionary Saving Unfettered," Working Paper.
- [11] Feigenbaum, James, (2008), "Information Shocks and Precautionary Saving," forthcoming in *Journal of Economic Dynamics and Control*.
- [12] Feigenbaum, James and Geng Li, (2008), "Lifecycle Dynamics of Income Uncertainty and Consumption," Working Paper.
- [13] Fernández-Villaverde, Jesús and Dirk Krueger, (2007), "Consumption over the Life Cycle: Some Facts from Consumer Expenditure Survey Data," *Review of Economics and Statistics* 89: 552-565.
- [14] Fogli, Alessandra and Fabrizio Perri, (2006). "The Great Moderation and the U.S. External Imbalance," *Monetary and Economic Studies* 26: 209-225.
- [15] Gordon, Roger H., (1984), *Differences in Earnings and Ability* (Garland: New York).
- [16] Gottschalk, Peter and Robert Moffitt, (1994), "The Growth of Earnings Instability in the U.S. Labor Market," *Brookings Papers on Economic Activity* 217-272.
- [17] Gourinchas, Pierre-Olivier and Jonathan A. Parker, (2002), "Consumption over the Life Cycle," *Econometrica* 70: 47-89.
- [18] Guvenen, Fatih, (2007), "Learning Your Earning: Are Labor Income Shocks Really Very Persistent?" *American Economic Review* 97: 687-712.
- [19] Huggett, Mark, (1996), "Wealth Distribution in Life-Cycle Economies," *Journal of Monetary Economics* 38: 469-494.
- [20] Jaimovich, Nir and Henry E. Siu, (2008), "The Young, the Old, and the Restless: Demographics and Business Cycle Volatility," NBER Working Paper 14063.
- [21] Lillard, Lee, and Yoram Weiss, (1979), "Components of Variation in Panel Earnings Data: American Scientists 1960-70," *Econometrica* 47: 437-454.
- [22] Kim, Chang-Jin and Charles Nelson, (1999), "Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle," *Review of Economics and Statistics* 81: 608-616.

- [23] McConnell, Margaret and Perez-Quiros, Gabriel, (2000), "Output Fluctuations in the U.S.: What Has Changed Since the Early 1980s?" *American Economic Review* 90 : 1464-1476.
- [24] Viceira, Luis M., (2001), "Optimal Portfolio Choice for Long-Horizon Investors with Nontradable Labor Income," *Journal of Finance* 56: 433-470.