Departament d'Economia Aplicada

Subjective Income Expectations, Canonical Models and Income Risk

Xavier Ramos
Christian Schluter

DOCUMENT DE TREBALL

03.10

Universitat Autònoma de Barcelona
Facultat de Ciències Econòmiques i Empresarials

Subjective Income Expectations, Canonical Models, and Income Risk

Xavi Ramos[*]
Universitat Autonoma de Barcelona

Christian Schluter[†]
University of Southampton

November 2003

Abstract

Expectations are central to behaviour. Despite the existence of subjective expectations data, the standard approach is to ignore these, to hypothecate a model of behaviour and to infer expectations from realisations. In the context of income models, we reveal the informational gain obtained from using both a canonical model and subjective expectations data. We propose a test for this informational gain, and illustrate our approach with an application to the problem of measuring income risk.

Keywords: subjective expectation data, canonical income models, income risk.

[*]Dept. Economia Aplicada, Universitat Autonoma de Barcelona, 08193 Bellaterra, Spain. Tel +34.93.5812868, Fax +34.93.5812292. Email: Xavi.Ramos@uab.es.

[†]Corresponding author. Department of Economics, University of Southampton, Highfield, Southampton, SO17 1BJ, UK. Tel. +44 (0)2380 59 5909, Fax. +44 (0)2380 59 3858. Email: C.Schluter@soton.ac.uk. http://www.economics.soton.ac.uk/staff/schluter/

1 Introduction

Expectations are central to behaviour. In the absence of explicit expectations data, the standard approach is to hypothecate a model of behaviour and to infer expectations from realisations. If the investigating econometrician does not observe all relevant information constituting the information set of the agent, such expectations are limited information rather than full information expectations. However, explicit expectations data are often available, as standard surveys often ask interviewees regularly about their (subjective) expectations concerning, for instance, income changes. Such data are typically ignored in empirical analyses. This paper investigates the informational gain from using subjective expectations data.

We examine this issue in the context of income models and the measurement of income risk. Variants of error-component models, based on the distinction between permanent and transitory components of income, have become the canonical model family in the literature to analyse the dynamics of individual income.[1] A related issue is income risk and its measurement. The extent of income risk informs decisions of (risk averse) agents, and directly influences their level of welfare. Policy evaluations rely on such measures. In the standard approach income risk is inferred from the residuals obtained from the canonical model. If the model, however, does not adequately describe the income generating process, or if the information sets of agents and econometrician differ substantially, then these residuals will only yield biased estimates of income risk. For instance, suppose an agent voluntarily plans to reduce labour supply in the next period, but that this intention is not recorded in the survey. The ensuing income fluctuation is therefore perfectly predictable by the agent, though not by the econometrician who would mistakenly interpret the fluctuation as a manifestation of income risk.

In order to assess the potential informational gain from using subjective expectations data on income changes, we use a following sequential procedure, which augments the standard approach of estimating the canonical error components model by one additional estimation step. As usual, the first estimation step consists of regressing, using OLS, realised income changes on a set of covariates, and obtain the fitted residuals. In the new second step, we use OLS to regress the fitted residuals against the subjective expectations, and obtain the second set of fitted residuals, which are then used to estimate the error-component structure in the usual way.[2] The issue is whether the coefficient of the subjective expectations is significantly different from zero. This benchmark value of zero obtains from the standard assumption made in the literature - our null hypothesis to be tested - that the canonical model does represent the income generating process, and that the agents use this model to make their point forecasts. The residuals from the first regression are then orthogonal to the subjective expectations variable, resulting in a coefficient of zero. Significant differ-

[1] See, for instance, with reference to U.S. data Lillard and Willis (1978), MaCurdy (1982), Abowd and Card (1989), Gottschalk and Moffitt (1995), and Baker (1997). A recent application to British data is Ramos (2003). A principal focus of this literature is the extent to which income inequality is permanent or transitory.

[2] The reasons for using a sequential rather than a simultaneous procedure will be explained below.

ences from zero demonstrate (orthogonal) information held by the agent which is not captured by the model. Moreover, in the presence of a significant coefficient, using the residuals from the first rather than the second regression yields an over-estimate of income risk faced by the agents.

Interest in subjective expectations data is not new. Maddala (1994) provides a fairly recent review. Newer work describes and analyses the validity of income or earnings expectations data (Dominitz (1998), Dominitz and Manski (1997), Das and van Soest (1997)), and uses these data to analyse income risk (Guiso et al. (2001), and to test for models adopted in the saving, portfolio choice, and consumption literature (Guiso et al. (1992, 1996)). In the context of consumption models one of the main interests is to tests for rational expectations, which are typically rejected empirically (Dominitz (2001), Das and van Soest (1999, 2001), Pistaferri (2000)). However, our concern is different. Rather than looking for a conflict between canonical (income) models and subjective expectations data, we examine the complementarity between the two. In particular, we investigate whether the expectations data contain information orthogonal to the canonical model, and whether this potential information gain can be usefully exploited. In our application to the problem of income risk, we illustrate how both data sources can be used in conjunction with each other, rather than discarding one in favour of the other.

The plan of the paper is as follows. Section 2 describes the data, focussing on the subjective expectations data, as well as two variates measuring actual realisations. We assess the information content of the subjective expectations data by comparing them to next periods realisations, and conclude that these data are informative. Section 3 sets out the Canonical Model, and Section 3.1 describes our way of incorporating the subjective expectations data in the Canonical Model. We then propose our test of the merit of this procedure. We turn to the empirical results in Section 3.2, whilst Sections 3.3 and 3.4 consider the implication in the context of variance regressions and measuring income risk. Section 4 concludes.

2 The Data

Our empirical analysis is based on the first ten waves of the British Household Panel Survey (BHPS), a popular longitudinal dataset built on similar principles as the Panel Study of Income Dynamics (PSID) in the U.S..[3] The income concept is dictated by the survey question about subjective income expectations, and refers to total (real) household income (and not personal earnings).[4] Our unit of analysis, however,

[3]The BHPS is a longitudinal panel data set consisting of some 5500 households (approximately 10000 individuals) first interviewed in the autumn of 1991 (wave 1) followed and re-interviewed every year subsequently. The initial sample represents a response rate of about 69% (proxies included) of the effective sample size. Wave-on-wave attrition rates for the subsequent waves are low. For a detailed discussion of BHPS methodology and representativeness see Taylor (1995), and http://iserwww.essex.ac.uk/bhps/.

[4]Our income measure is the log of the pre-tax post-transfer real household income. It includes earnings from employment and self-employment, cash social security and social assistance benefits, and income from savings and investment. See Bardasi, Jenkins and Rigg (1999) for more detailed

is the person, since households can form and dissolve over time. We control for demographic characteristics directly rather than using an equivalisation procedure. We focus on households whose heads are "prime-aged" (between 25 and 59 years), and thus exclude "young" and "older" households because their life-cycle and in particular labour markets events can differ fundamentally (the "getting started" phenomenon, and living in retirement). Table 7 in the Appendix provides summary statistics of the sample.

We proceed to describe the income and subjective expectations data in some detail. Expectations are reported at time t while data about actual income changes are obtained from the $t+1$ wave. The subjective expectation data at our disposal are answers to the following survey question: "Looking ahead, how do you think you will be financially a year from now, will you be: better off, worse off than you are now, or about the same?" The data is thus categorical[5] rather than in a probabilistic format which would allow the estimation of the entire subjective probability distribution (as in e.g. Dominitz and Manski (1997) or in Guiso et al. (1992)).

We use two sources for information about actual, realised income changes. First, each wave of the survey includes as question, phrased on the same lines as the expectations questions, about actual income changes experienced in this period ("looking back ..."). Although the income concept in these survey questions is not explicitly defined, it is highly likely that the respondent will use the same metric in his answers. As for the expectations data, the answers are categorical. We refer to this data as "perceived income changes".

We also use estimates of total household income generated by the data providers (which we will refer to as "actual incomes"). This realisation data is continuous, but the income concept of the data provider might not fully correspond to that of the respondent. However, both income concepts are highly, though not perfectly, correlated. Figure 1 plots the time series of the mean of the actual income changes for the three categories of perceived actual income changes. The time-averaged means of actual income change equal roughly -2% (perception: worse off), 3% (about same), and 6% (better off).[6]

Table 1 reveals the information content of the subjective expectations data, by juxtaposing expectations with the perceived actual income changes reported by respondents one year after the former. The expectations data do have information content, as the matrix differs significantly from one with all entries equal to 1/3. On average, expectations are fulfilled in terms of the given categories, as, for instance, 51% of individuals expecting no change actually report no change in the next period. However, a large proportion of individuals also make wrong forecast. For instance, 22% of people, who expected to be worse off, reported in the next period to have

information on the BHPS income variable.

[5]For an examination of the potentially different interpretations of such categorical responses, see Manski (1990).

[6]Other ways of estimating the relationship between perceived and actual income changes provide very similar results. A random effects regression of actual income change on subjective expectations and time dummies (to capture macro shocks) provides the following estimates for the means of actual income change: -1.6% (worse off), 3% (about same), and 5.2% (better off).

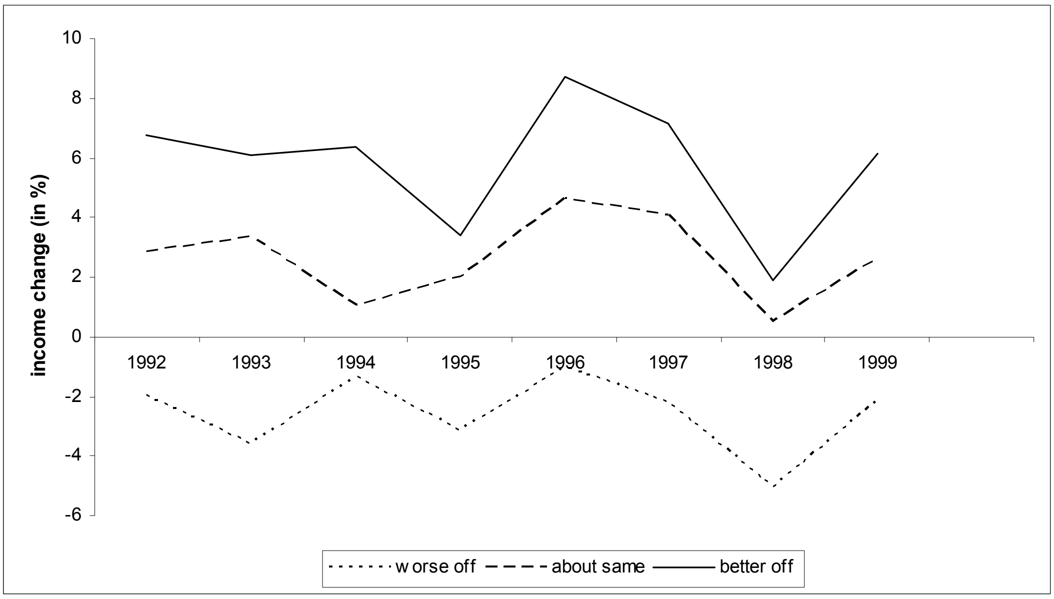

Figure 1: "Perceived income changes" and "actual income changes": Mean of actual income change by subjective perceptions.

actually become better off.

The table also disaggregates the sample according to educational status in terms of school-leaving qualifications. The category 'O levels or less' refers to individuals who have left school aged 16 or younger, and 'A-levels' to school leavers who have graduated with a university entrance qualification (school grade 12) . Interestingly, the more highly educated individuals do not hold, on average, more accurate expectations. Many corresponding cells are not statistically significant different, and educational covariates included in the types of regressions reported below turn out to be statistically insignificant. We have therefore desisted from partitioning the sample by educational category in the remainder of the analysis.

This expectations data for Britain share many features of expectations data reported for other countries. For instance, using probabilistic expectations data for the U.S., Dominitz (1998) finds that (i) expectations of university educated respondents were not more accurate than those of individuals with lower education, and (ii) many respondents were too optimistic ex post. This is also true for our sample as 29.51% of the population were too optimistic, compared to 20.9% who were too pessimistic (49.58% of individuals got it right). In the Netherlands, however, pessimists appear outnumber the optimists (Das and van Soest, 1999).

Table 2 quantifies the extent of the relative expectations error by reporting the mean of actual income growth for the various expectations and perception groups. The last row are the time-averaged means for Figure 1. All rows are, reassuringly, monotonic, as are the row and column for the totals. We note, however, that the group expecting be worse off actually experienced a small income improvement (1.3%, but

expectations\	perceived actual income change			
	worse off [%]	about same [%]	better off [%]	Total [%]
		Entire sample		
worse than now	44.26	33.99	21.75	100
about the same	22.16	50.98	26.86	100
better than now	18.03	32.86	49.11	100
		University or higher degree		
worse than now	37.71	33.14	29.14	100
about the same	20.49	44.72	34.79	100
better than now	16.83	24.32	58.85	100
		A Levels or higher vocational degree		
worse than now	47.52	30.20	22.28	100
about the same	23.25	47.32	29.44	100
better than now	18.59	31.26	50.15	100
		O Levels or less		
worse than now	43.93	37.96	18.11	100
about the same	21.68	55.51	22.81	100
better than now	17.96	38.40	43.64	100

Table 1: Subjective expectations and "perceived income changes."

expectations\	perceived income change			
	worse off	about same	better off	Total
worse than now	-3.0	2.5	7.9	1.3
about the same	-1.4	2.7	5.3	2.5
better than now	-3.6	4.1	7.1	4.2
Total	-2.3	3.0	6.4	2.9

Table 2: Mean of "actual income changes" (in per cent) by subjective expectations and "perceived income changes."

recall that "actual" and "perceived" income changes are highly but not perfectly correlated), but reassuringly the income improvements for the other two expectations groups are monotonically increasing.

Finally, Table 3 describes who holds what expectations. We limit our discussion to aspects relevant in the context of the canonical income models, explained in detail in the next section (as a deep analysis of expectation formation is not our goal in this paper).[7] Reported are the estimates of ordered logits of the three expectations categories. The baseline specification, labelled I, includes a set of standard covariates measuring demographics and labour market characteristics (including dummies for: self-employment, part and full time worker, whether the job is permanent or temporary ("contract"), and the standard occupational classification). We have included occupational status as indicators of the potential complexity of the income

[7]Summary statistics are shown in Appendix Table 7.

generating process, and these turn out to be significant. Many level variables, in particular income levels and time-invariant covariates such as education, are not significant, whereas many "event" variables, referring to events between $t-1$ and t, are significant.

We also report an augmented specification, labelled II, which adds to I indicators for the perceived income change, and the past expectations error.[8] The significance of all these variables suggests a dynamic structure of how expectations are formed. Perceived income changes in the past twelve months seem to have an impact on expectation formation. People who suffered a subjective income fall (gain) are more (less) pessimistic about their future income than people who have experienced no income change. This phenomenon has also been observed for the Netherlands (Das and van Soest, 1999).

Individuals do not seem to revise their expectations in the light of past prediction errors. Individuals who underestimated their incomes in the past are more likely, and people who overestimated them are less likely to get their predictions wrong than people who saw their expectations fulfilled; i.e. the predictions of the underpredictors do not, on average, improve. Moreover, people who "underpredicted a lot"are more pessimistic than those who "underpredicted a little", while individuals who "overpredicted a lot" are less pessimistic than those who "overpredicted a little".

	I		**II**	
	Coef.	P-value	Coef.	P-value
Age	-0.087	0.000	-0.055	0.000
Age squared	0.001	0.001	0.0004	0.031
Sex: Female	-0.201	0.000	-0.170	0.000
Education	included	not significant	included	not significant
Health limits work: Yes	-0.145	0.005	-0.119	0.025
Marital Status: Married	0.035	0.341	-0.014	0.715
Children: 1 child	-0.123	0.001	-0.068	0.087
more than 1 child	0.020	0.582	0.043	0.253
decreased	0.023	0.723	0.065	0.330
increased	0.004	0.958	0.387	0.000
Job Status: self-empl.	0.326	0.000	0.236	0.000
Became self-employed	0.215	0.048	0.261	0.023
Became employee	0.378	0.001	0.273	0.021
Full-time or PT: Part-time	0.025	0.535	0.049	0.241
Became Part-time	0.062	0.428	0.117	0.165
Became Full-Time	0.116	0.105	0.002	0.984

[8]This variable is defined as the deviation between expectations and perceived income change."Underpredict a lot" refers to individuals holding "worse than now" expectations and reporting being "better off" in the perceived income change question, whereas "underpredict a little" refers to individuals holding "worse than now" ("about the same") expectations and reporting being "about the same" ("better off") in the perceived income change question (see Table 2).

	I		II	
	Coef.	P-value	Coef.	P-value
Contracts: Non-standard	0.176	0.027	0.137	0.103
Became non-standard	-0.109	0.359	-0.088	0.487
Became Permanent	0.241	0.001	0.129	0.110
Stan.Occupation Class.				
Professional	-0.147	0.005	-0.114	0.035
Associate profess. technical	-0.112	0.027	-0.077	0.139
Clerical	-0.117	0.016	-0.066	0.187
Craft related	-0.289	0.000	-0.189	0.000
Protective services	-0.168	0.002	-0.099	0.079
Sales	-0.041	0.518	-0.026	0.696
Plant machine operatives	-0.328	0.000	-0.239	0.000
Other	-0.320	0.000	-0.227	0.001
Got a "worse" occupation	0.210	0.000	0.157	0.000
Got a "better" occupation	0.149	0.000	0.103	0.019
Region: North	-0.045	0.082	-0.018	0.491
Changes region	0.497	0.000	0.530	0.000
Time Dummies	included	significant	included	significant
Log (income level)	included	not significant	included	not significant
Δ in income$_{t-1,t}$: "worse off"			-0.966	0.000
"better off"			1.547	0.000
Expectations Error$_{t-1,t}$ *				
"underpredict a lot"			-1.505	0.000
"underpredict a little"			-0.912	0.000
"overpredict a little"			1.002	0.000
"overpredict a lot"			2.198	0.000
Number of obs		25277		24531
P-value of χ^2		0		0
Log likelihood		-22705.2		-20687.3

Table 3: Ordered logits: Who holds what expectations ? Notes: * Variable defined in Footnote 8.

3 The Canonical Income Model and Subjective Expectations

The Canonical Model for income[9] is based on the standard distinction between permanent and transitory income. The permanent component typically takes a Mincerian form (including polynomials in age), and changes in this component are often modelled by a random walk specification. Unobserved ability or heterogeneity gives rise to

[9] See Footnote 1 for a selected list of references.

a fixed effect. In empirical analyses, the transitory income component usually follows an ARMA(1,1) or MA(1) specification (with, of course, different dynamic implications). Since the standard specification test leads us to adopt a MA(1) specification in our empirical application (see section 3.2.3 for details), we confine our exposition to this case.

To be precise, the Canonical Model takes the form

$$y_{it} = \beta' Z_{it} + u_{it},$$

where y_{it} denotes the log of income of person i at time t, Z_{it} is a matrix of time-varying and time invariant covariates, and u_{it} denotes the error term. We consider the following standard specification of the error-component structure

$$\begin{aligned} u_{it} &= \alpha_i + p_{it} + \varepsilon_{it} \\ p_{it} &= p_{it-1} + \eta_{it} \text{ with } \eta_{it} \sim iid\left(0, \sigma_\eta^2\right) \\ \varepsilon_{it} &= \nu_{it} - \delta\nu_{it-1} \text{ with } \nu_{it} \sim iid\left(0, \sigma_\nu^2\right),\ \varepsilon_{i0} \sim iid\left(0, \sigma_{\varepsilon_0}^2\right) \end{aligned} \tag{1}$$

where α_i represents a fixed effect, p_{it} a permanent income component which follows a random walk with stationary innovations, and ε_{it} the transitory income component which evolves as a MA(1) process. $\sigma_{\varepsilon_0}^2$ denotes the variance of the initial condition. It is customary to difference out the fixed effects by examining income growth

$$\begin{aligned} \Delta y_{it} &= y_{it} - y_{it-1} \\ &= \beta' \Delta Z_{it} + r_{1,it}. \end{aligned} \tag{2}$$

Income growth residuals $\widehat{r_{1,it}} = \Delta u_{it}$ can then be expressed as

$$\widehat{r_{1,it}} = \eta_{it} + \nu_{it} - (1+\delta)\,\nu_{it-1} + \delta\nu_{it-2}, \tag{3}$$

implying the covariance structure[10]

$$E\left\{\widehat{r_{1,it}}\widehat{r_{1,it-s}}\right\} = \begin{cases} E\left\{\eta_{it}^2\right\} + E\left\{\nu_{it}^2\right\} + (1+\delta)^2 E\left\{\nu_{it-1}^2\right\} + \delta^2 E\left\{\nu_{it-2}^2\right\} & \text{for } s=0 \\ -(1+\delta)\,E\left\{\nu_{it-1}^2\right\} - \delta(1+\delta)\,E\left\{\nu_{it-2}^2\right\} & \text{for } s=1 \\ \delta E\left\{\nu_{it-2}^2\right\} & \text{for } s=2 \\ 0 & \text{for } s>2. \end{cases} \tag{4}$$

The typical estimation procedure for the Canonical Model consists of two stages. In the first stage, (2) is estimated using OLS. This yields the fitted residuals $\widehat{r_{1,it}}$. In the next stage, the error-component structure is recovered from the empirical covariance matrix of $\widehat{r_{1,it}}$ using (4) by minimum distance methods.

[10]Note that due to the initial condition, the variance at $t = 0$, $E\left\{\widehat{r_{1,i0}}\widehat{r_{1,i0}}\right\} = E\left\{\eta_{it}^2\right\} + E\left\{\varepsilon_{i0}^2\right\}$,differs from (2).

3.1 Incorporating Subjective Expectations

It is customary in the literature to hypothecate that the Canonical Model correctly specifies the income generating process, and that the agents' use this model to make income forecasts (i.e. agents have "rational expectations"). However, it is conceivable that the econometrician does not observe all relevant information at the disposal of agents', so that his limited information expectations differ from the agents' full information expectations. Other possibilities are that agents' do not use the Canonical Model to make forecasts, or that the Canonical Model misspecifies the income generating process. Using subjective expectations data of the form described in the Data Section, these issues can be tested empirically.

In order to assess the information gain from using subjective expectations data on income changes we augment the standard approach of estimating the Canonical Model by one additional estimation step. The first estimation step is the standard one described above. The next estimation step is new, and consists of the OLS regression of the fitted residuals on the subjective expectations data, denoted by E_{it}

$$\widehat{r_{1,it}} = \gamma E_{it} + r_{2,it}. \tag{5}$$

This is our key regression.[11] Denote the fitted residuals of this second stage regression by $\widehat{r_{2,it}}$. The third estimation step is again standard, and recovers the error-component structure from $\widehat{r_{2,it}}$ in the usual way using minimum-distance techniques. To be precise, the error component structure for the MA(1) case is now given , in a slight abuse of notation, by

$$\widehat{r_{2,it}} = \eta_{it} + \nu_{it} - (1+\delta)\,\nu_{it-1} + \delta\nu_{it-2}, \tag{6}$$

and the relevant covariance structure is $E\left\{\widehat{r_{2,it}}\widehat{r_{2,it-s}}\right\}$. We refer to this specification as the expectations augmented model. Unless $\gamma = 0$, the specifications (3) and (6) differ, and the Canonical Model is misspecified.

3.1.1 The Test: Does incorporating subjective expectations yield informational gain ?

The principal object of our interested is the coefficient γ in the stage two regression (5). The issue is whether this coefficient of the subjective expectations is significantly different from the benchmark value of zero,

$$H_0: \ \gamma = 0.$$

This null hypothesis obtains under the assumption that the Canonical Model does represent the income generating process, and that the agents use this model to make

[11] Rather than using E_{it} as an extra regressor in the standard first step, our estimation procedure is sequential for the following reason. Assume that the canonical model does perfectly describe the income generating process, that agents use the model to make predictions, and that the econometrician observes all relevant variables so that everyone's expectations are full information expectations. In these circumstances, perfect collinearity between the model and E_{it} would make impossible the estimation of γ in a single augmented regression, but creates no problem for the sequential procedure (in which case the benchmark value of $\gamma = 0$ obtains).

their point forecasts. Deviations from the model-based predictions are then only the result of surprises, and the residuals from the first regression orthogonal to the subjective expectations variable. In these circumstances, the expectation data contains no additional information, and hence $\gamma = 0$. On the other hand, a significant difference of the coefficient estimate from zero implies that either agents hold (orthogonal) information which is not captured by the model, or they do not use the Canonical Model to make point forecasts.

3.2 Empirical Results

We proceed to discuss the empirical results based on the BHPS, taking each estimation stage in turn. Since our object of interest is γ, our choice of regressors for the first stage regression is determined by current standard practice: the empirical implementation of the Canonical Model is also canonical. Also, the results from the conventional stages one and three will only be discussed very briefly. We mainly note that these are consistent with the findings in the literature.

3.2.1 The Conventional Stage 1 Regression

Our choice of regressor ΔZ_{it} in the stage one regression (2) includes all covariates listed for column I in Table 3. Detailed results, not reported here for the sake of brevity, are available from the authors on request. Our set of regressors extends the canonical one by including covariates which are typically fixed for long periods if not time invariant, such as occupation. This provides an additional check on the Canonical Model, since the coefficients of these covariates should be zero. As it turns out, this is typically the case: in particular, the regressors relating to education and occupation are all insignificant.

3.2.2 The Key Stage 2 Regression

The results of the key OLS estimation of equation (5) are reported in Table 4. It is evident that the γ coefficients are significantly different from zero. We therefore conclude that the subjective expectations data do contain information which is not incorporated in the Canonical Model, and the null H_0 has to be rejected. This is one of our key empirical findings. We discuss the quantitative and economic aspects of this rejection in the next section in the context of measuring income risk.

	$\widehat{\gamma}_{OLS}$	SE
expectations:		
"about the same"	0.028	(0.007)
"better off"	0.047	(0.008)
constant	-0.031	(0.006)
R^2	0.0016	

Table 4: Estimation of γ in the key second stage regression.

	Canonical Model ($\gamma = 0$)		expectations augmented model ($\gamma = \widehat{\gamma}_{OLS}$)	
σ_η^2	0.0519	(10.4)	0.0517	(10.3)
$\sigma_{\varepsilon_0}^2$	0.0724	(12.2)	0.0722	(12.2)
σ_ν^2	0.0421	(12.3)	0.0420	(12.3)
δ	-0.2645	(-8.7)	-0.2629	(-8.7)
SSR , $\chi^2 - test$	0.0051	72.215	0.0051	72.748

Table 5: Estimates of the error component structure. t-ratios in parenthesis.

3.2.3 The Conventional Stage 3 Regression

We recover the error-component structure from the stage 2 residuals $\widehat{r_{2,it}}$ in the standard way. Our tests, details of which are given in the Appendix, lead us to adopt (i) a MA(1) specification for the transitory component, (ii) reject the absence of permanent shocks. In short, we model log earnings y_{it} as a random walk plus a $MA(1)$ process, which implies that the second stage regression residuals $\widehat{r_{2,it}}$ follow the specification described by equation (6).

The estimates of the error component parameters are reported in Table 5. As a benchmark, we have also estimated the Canonical Model, with the MA(1) specification given by (3). Comparing the results for both models, it turns out that the estimates of the error-component parameters are statistically not significantly different. Thus, despite the significance of $\widehat{\gamma}$ in the second stage regression, ignoring the informational gain afforded by the subjective expectations has no significant impact on the error component estimates.

We briefly turn to the interpretation of the error component estimates for the expectations augmented model (6), which result in the following estimated covariance structure

$$E\left\{\widehat{r_{2,it}}\widehat{r_{2,it-s}}\right\} = \left\{\begin{array}{ll|ll} 0.1239 & \text{for } s=0, t=0 & 0.1195 & \text{for } s=0, t>0 \\ -0.02728 & \text{for } s=1, t=0 & -0.02282 & \text{for } s=1, t>0 \\ -0.00658 & \text{for } s=2, t=0 & -0.01104 & \text{for } s=2, t>0 \\ 0 & \text{for } s>2, t=0 & 0 & \text{for } s>2, t>0. \end{array}\right.$$

The predicted variance is 0.124 for the initial period and 0.120 for the other 8 periods (the difference being due to the initial condition). The predicted permanent component is 0.052 and the predicted transitory one is 0.072 for the first period and 0.068 for the other periods (Table 5). That is, the permanent component accounts for 43 per cent and transitory component for 57 per cent of the variance of the income growth residuals. The $MA(1)$ process implies that there is no persistent serial correlation in the (unexplained) income growth rates. The negative estimate of the MA parameter δ accommodates the (unreported) sharp decline in the covariances at the first and second orders. Our results for the decomposition into permanent and transitory components are similar to those obtained in other similar studies, as are the estimates of the other error component parameters.[12]

[12]It is not possible to compare directly our results with previous studies as these differ in terms

3.3 Implications for Variance Regressions

Apart from the conditional mean, the conditional variance plays an important role in its own right in some applications. The leading example is the literature on precautionary savings, in which the conditional variance determines consumption and savings. We therefore turn to the implications of incorporating informative subjective expectations for variance regressions.[13]

Denote the information set of agent i at time t by Ω_{it}. If, under the null hypothesis, agents use the Canonical Model, then substracting from (2) its conditional expectation, and squaring, yields

$$Var\left\{\Delta y_{it}|\Omega_{it-1}\right\} = Var\left\{r_{1,it}|\Omega_{it-1}\right\}.$$

Hence conditional variances in the Canonical model can be estimated from the empirical variance of the fitted stage-one residuals. In the MA(1) model we have

$$Var\left\{\widehat{r}_{1,it}|\Omega_{it-1}\right\} = \sigma^2_{\eta,c} + \sigma^2_{\nu,c}, \tag{7}$$

(where the subscript c refers to the Canonical model, and e will refer below to the expectations augmented model). However, if agents' subjective expectations contain information orthogonal to the canonical model, i.e. if γ is non-zero, then stage-two residuals should be used, and the relevant variance is the conditional full information variance of the expectations augmented model, namely

$$Var\left\{\widehat{r}_{2,it}|\Omega_{it-1}\right\} = \sigma^2_{\eta,e} + \sigma^2_{\nu,e}. \tag{8}$$

How large is the discrepancy between the two variances for our data ? Based on the results reported in Table 5 we obtain the following estimates: $\widehat{Var}\left\{\widehat{r}_{1,it}|\Omega_{it-1}\right\} =$ 0.0940, and $\widehat{Var}\left\{\widehat{r}_{2,it}|\Omega_{it-1}\right\}$ =0.0937. Ignoring the informational gain contained in the subjective expectations has only a small effect. This is a direct consequence of the results reported in the last section: despite the rejection of $\gamma = 0$, Canonical and expectations augmented models yield similar error component estimates.

3.4 Implications for the Measurement of Income Risk

Beside the literature on precautionary savings and variance regressions, a separate literature on the measurement of income risk has recently emerged, since outcomes other than consumption and savings might be of interest. As it turns out, the results of the previous section can be easily imported.

---

of income conepts, and country focus. This notwithstanding, our parameter estimates are very plausible when compared to PSID-based studies (see, *inter alia*, Meghir and Pistaferri (2003), Baker (1997), Carroll and Samwick (1997), MaCurdy (1982)). The study most comparable to our set up is Carroll and Samwick (1997) for the U.S.. They find a more unequal relative contribution of the permanent and transitory components to total income growth variance (33 and 77 per cent, respectively).

[13]For a further discussion of variance regressions see Dominitz (2001). Meghir and Pistaferri (2003) propose to impose an ARCH specification on the conditional variance.

One popular measure of income risk is the (unconditional) variance of the income growth residuals (Burgess et al., 2000 and references therein), given by $Var\{\widehat{r}_{1,it}\}$, or the coefficient of variation (as in Guiso et al., 2001). For the Canonical Model with the MA(1), this unconditional variance equals

$$Var\{\widehat{r}_{1,it}\} = 2\sigma^2_{\nu,c}\left(1 + \delta_c + \delta_c^2\right) + \sigma^2_{\eta,c},$$

which we also refer to as the limited information variance. However, it is conventionally assumed that agents make full use of their information at their disposal, in particular their knowledge of the dynamics of the income process. If agents use the Canonical Model to make their predictions, the appropriate measure of income risk is the full information variance

$$risk_c = Var\{\widehat{r}_{1,it}|\Omega_{it-1}\}, \tag{9}$$

with $risk_c \leq Var\{\widehat{r}_{1,it}\}$. If the agents' subjective expectations contain information orthogonal to the canonical model, i.e. if γ is non-zero, then the full information variance of the expectations augmented model is the appropriate measure of risk:

$$risk_e = Var\{\widehat{r}_{2,it}|\Omega_{it-1}\}. \tag{10}$$

For the MA(1) model, these variances are given in (7) and (8).

We turn to the quantifications of these measures. Based on the results reported in Table 5 we obtain the following estimates: $\widehat{Var}\{\widehat{r}_{1,it}\} = 0.1202$, $\widehat{risk}_c = 0.0940$, and $\widehat{risk}_e$=0.0937. Ignoring the agents knowledge of the income dynamics governed by the Canonical Model leads to overestimate income risk by a dramatic 28%. Once the full information variance is adopted as the appropriate measure of risk, the informational gain from incorporating subjective expectations is small.

4 Conclusion

Given the popularity of the Canonical Model of income, one important issue is whether agents make use of this model when forming expectations (since it is these expectations which inform the decisions of agents). We have investigated this question using subjective expectations data elicited from British households. Rather than seeking to reject one approach in favour of the other, we have focused on the complementarity between the two: we assess the informational gain from using expectations data by augmenting the Canonical Model. This new intermediate regression consists of regressing stage one residuals on the expectations data. Our first empirical finding is that the coefficients of the expectations data are significant. We have assessed the economic importance of this by juxtaposing the Canonical Model and the expectations augmented model in terms of the estimates of the error component structure, variance regressions, and measures of income risk. It turns out that the differences between the two are only small, which vindicates the use of the Canonical Model in empirical work.

References

Abowd, J. and D. Card (1989), "On the covariance structure of earnings and hours changes", *Econometrica*, 57, 411-445.

Baker, M. (1997), "Growth-rate heterogeneity and the covariance structure of life-cycle earnings", *Journal of Labour Economics*, 15 (2), 338-75.

Baker, M. and G. Solon (1997), "Earnings dynamics and inequality among Canadian men, 1976-1992: evidence from longitudinal income tax records". NBER Working Paper no. 7370.

Bardasi, E., S. P. Jenkins and J. A. Rigg (1999), "Documentation for derived current and annual net household income variables, BHPS waves 1-10", Institute for Social and Economic Rerearch, University of Essex.

Burgess, S., K. Gardiner, S. P. Jenkins and C. Propper (2000), "Measuring income risk", CEPR Discussion Paper: 2512.

Cappellari, L. (2000), "The dynamics and inequality of Italian male earnings: permanent changes or transitory fluctuations?", ISER Working Paper no. 2000-41, University of Essex: Colchester.

Carroll, C. D. and A. A. Samwick (1997), "The nature of precautionary wealth", *Journal of Monetary Economics,* 40, 41-71.

Das, M. and A. van Soest (1997) "Expected and realized income changes: Evidence from the Dutch socio-economic panel", *Journal of Economic Behavior and Organization*, 32, 137-154.

Das, M. and A. van Soest (1999) "A panel data model for subjective information on household income growth", *Journal of Economic Behavior and Organization*, 40, 409-426.

Das, M. and A. van Soest (2001) "Expected versus realized income changes: A test of the rational expectations hypothesis", Tilburg University.

Dickens, R. (2000) "The evolution of individual male wages in Great Britain: 1975-95", *Economic Journal*, 110, 27-49.

Dominitz, J. (1998) "Earnings expectations, revisions, and realisations", *The Review of Economics and Statistics*, 80, 374-388.

Dominitz, J. (2001), "Estimation of income expectations models using expectations and realization data", *Journal of Econometrics*, 102, 165-195.

Dominitz, J. and C. F. Manski (1997), "Using Expectations Data To Study Subjective Income Expectations", *Journal of the American Statistical Association*, 92, No. 439, 855-867.

Gottschalk, P. and R. Moffitt (1995), "Trends in the covariance structure of earnings in the U.S.: 1969-1987", Working Paper, Brown University.

Guiso, L., Jappelli, T., and L. Pistaferri (2001), "An empirical analysis of earnings and employment risk", *Journal Business and Economics Statistics,* 20(2), 242-254.

Guiso, L., Jappelli, T., and D. Terlizzese (1992), "Earnings uncertainty and precautionary saving", *Journal of Monetary Economics,* 30, 307-337.

Guiso, L., Jappelli, T., and D. Terlizzese (1996), "Income Risk, Borrowing Constraints, and Portfolio Choice", *American Economic Review,* 86, 158-172.
MaCurdy, T. (1982), "The use of time series processes to model the error structure of earnings in longitudinal data analysis", *Journal of Econometrics*, 18, 83-114.
Maddala, G.S. (1994), "Survey data on expectations: what have we learnt ?", in *Econometric Methods and Applications*, Vol. 1 ed. G.S. Maddala, Brookfield, VT: E. Elgar.
Manski, C. (1990), "The use of intentions data to predict behaviour: a best-case analysis", *Journal of the American Statistical Association*, 85, 934-940.
Meghir, C. and L. Pistaferri (2003), "Income variance dynamics and heterogeneity", *Econometrica*, forthcoming.
Pistaferri, L. (2001), "Superior Information, Income Shocks, and the Permanent Income Hypothesis", *The Review of Economics and Statistics*, 83, 465-476.
Ramos, X. (2003) "The covariance structure of earnings in Great Britain, 1991-1999", *Economica*, 70, 353-374.
Taylor, M.F. (ed.) (1995), *British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices*, Colchester.

A Error Component Specifications

The MA(1) specification has been adopted via the following specification search. We started out fitting a low order ARMA specification for the transitory component ε_{it}. However, the estimates of this specification suffered from some common identification problems.[14] Thus, we assumed that ε_{it} follows an $MA(q)$ process, with q to be determined empirically by testing the suitability of the theoretical structure imposed on the estimated autocovariances $E\left\{\widehat{r_{2,it}}\widehat{r_{2,it-s}}\right\}$. Specifically, the MA(1) process given by (6) imposes zero-restrictions on third and higher order covariances (see equation (4)). Table 6 reports the test results[15] of zero-restrictions for the null hypothesis that $E\left\{\widehat{r_{2,it}}\widehat{r_{2,it-s}}\right\} = 0$ (for $1 \leq s \leq 5$). Unexplained income growth rates (i.e. stage 2 residuals $\widehat{r_{2,it}}$) seem to be correlated up to the second order: the test statistic does not reject the null of zero autocovariances at third or higher order (p-value $= 0.122$), but rejects it at second or higher order (p-value < 0.001). Therefore we use an $MA(1)$ process to model the transitory component.

	Order of covariances				
	1	2	3	4	5
degrees of freedom	36	28	21	15	10
χ^2	401.3	111.0	28.7	9.9	7.2
p-value	<0.001	<0.001	0.12	0.82	0.71

Table 6: Tests of zero-restrictions for MA(q) order.

Next, we test for the absence of a permanent shock. Following Meghir and Pistaferri (2001) we do so by testing whether the variance of the permanent shock is zero. Given that the transitory component follows a $MA(1)$ process, the covariance structure of stage 2 residuals $r_{2,it}$ takes, under the null hypothesis of no permanent shock, the form

$$E\left\{r_{2,it}r_{2,it}\right\} = \begin{cases} E\left\{\nu_{it}^2\right\} + (1+\delta)^2 E\left\{\nu_{it-1}^2\right\} + \delta^2 E\left\{\nu_{it-2}^2\right\} & \text{for } s=0 \\ (1+\delta) E\left\{\nu_{it-1}^2\right\} - \delta(1+\delta) E\left\{\nu_{it-2}^2\right\} & \text{for } s=1 \\ \delta E\left\{\nu_{it-2}^2\right\} & \text{for } s=2 \\ 0 & \text{for } s>2. \end{cases} \tag{11}$$

Hence

$$E\left\{r_{2,it}\left(r_{2,it-2} + r_{2,it-1} + r_{2,it} + r_{2,it+1} + r_{2,it+2}\right)\right\}$$

[14] In particular, the estimated variance of the random walk σ_η^2 is negative. Such identification problems are not unusual in the literature, and have been encountered, inter alia, by Ramos (2003), Capellari (2000), Dickens (2000), Baker and Solon (1999) and Baker (1997).

[15] See e.g. Abowd and Card (1989). The test statistic has the form $m_r' V_{rr}^{-1} m_r$, where m_r is the subvector comprising the elements of the covariance matrix restricted to zero and V_{rr} is the covariance matrix associated to the elements of m_r. The test statistic is distributed as a χ^2 with degrees of freedom equal to the dimension of m_r.

equals zero under the null. Under the alternative hypothesis, however, it equals the variance of the permanent shock $E\{\eta_{it}^2\}$. Since we do not observe the true shocks to income, we use the predicted composite residuals from (5) to implement this test, which is a standard one-sided test. The null hypothesis of no permanent shock is strongly rejected by the data: The t-statistic is 3.45, which implies a p-value smaller than 0.01.

B Data Appendix

Table 7 provides summary statistics of the sample. The educational dummies refer to school-leaving qualifications. The category 'O levels or less' refers to individuals who have left school aged 16 or younger, and 'A-levels' to school leavers who have graduated with a university entrance qualification (school grade 12). The categories of the "expectations error" are defined in footnote 8.

	Mean	SD
Age	40.66	9.20
Sex: Female	0.47	0.50
Education (increasing in levels)		
No Qualifications	0.14	0.35
Other Qualifications	0.09	0.28
O levels	0.21	0.41
A levels	0.12	0.32
Other Higher Qualifications	0.29	0.46
Health limits work: Yes	0.07	0.26
Marital Status: Married	0.80	0.40
Children		
1 child	0.18	0.39
More than 1 child	0.26	0.44
Decreased number of children	0.04	0.21
Increased number of children	0.03	0.18
Job Status		
Self-employed	0.12	0.33
Became self-employed	0.02	0.13
Became employee	0.02	0.12
Full-time or Part-time		
Part-time	0.19	0.39
Became Part-time	0.03	0.17
Became Full-Time	0.03	0.18
Contract		
Non-standard	0.05	0.22
Became non-standard	0.02	0.15
Became Permanent	0.03	0.17
Region		
North	0.58	0.49
Does change region	0.02	0.12

	Mean	SD
Stan.Occupation Class.		
Professional	0.12	0.33
Associate profess. technical	0.12	0.33
Clerical	0.17	0.38
Craft related	0.12	0.33
Protective services	0.09	0.29
Sales	0.06	0.23
Plant machine operatives	0.09	0.28
Other	0.06	0.24
Got a "worse" occupation	0.10	0.30
Got a "better" occupation	0.11	0.32
Time Dummies		
Wave 2	0.12	0.33
Wave 3	0.11	0.31
Wave 4	0.11	0.31
Wave 5	0.11	0.31
Wave 6	0.11	0.31
Wave 7	0.11	0.32
Wave 8	0.11	0.31
Wave 9	0.11	0.31
Log (income level)	9.75	0.50
Δ in income$_{t-1,t}$		
"worse off"	0.23	0.42
"better off"	0.33	0.47
Expectations Error$_{t-1,t}$ *		
"underpredict a lot"	0.02	0.14
"underpredict a little"	0.18	0.39
"overpredict a little"	0.23	0.42
"overpredict a lot"	0.06	0.23

Table 7: Summary statistics for the pooled sample. N=24,531 Notes: * Variable defined in Footnote 8.

Aquest document pertany al Departament d'Economia Aplicada.
Data de publicació: *desembre **2003***

Departament d'Economia Aplicada
Edifici B
Campus de Bellaterra
08193 Bellaterra

Telèfon: (93) 581 1680
Fax: (93) 581 2292
E-mail: d.econ.aplicada@uab.es
Http: //www.uab.es/dep-economia-aplicada/