Costly Information Processing and Consumption Behavior*

Stefan Hoderlein
Boston College

Michael Peters
Yale

November 2012

Abstract

What information do individuals use when they form expectations about future events? In this paper we present an econometric framework to answer that question in a world where individuals are profoundly heterogeneous. We first show how individual information sets can be characterized by simple nonparametric exclusion restrictions and provide a quantile based test for costly information processing. We then use microdata on individual income expectations to study what information agents pay attention to when forecasting future earnings. Consistent with models where information processing is costly, we find that individuals' information sets are coarse in that valuable information is discarded. To quantify the welfare effects, we calibrate a standard consumption life-cycle model. Consumers would be willing to pay 0.035% of their permanent income to incorporate the econometrician's information set in their forecasts. This represents a lower bound on the costs of information processing.

1 Introduction

Individuals' expectations about uncertain events are a key aspect of modern economics. Knowing what expectations individuals hold is therefore crucial to understand and predict behavior Manski (2004). A key ingredient in the process of expectation formation is the information set agents employ. In this paper we estimate the content of information sets using micro data

*Stefan Hoderlein: Department of Economics, Boston College, eMail: stefan_hoderlein@yahoo.com. Michael Peters: Department of Economics, Yale University, eMail: m.peters@yale.edu. We have received many helpful comments from Daron Acemoglu, Marios Angeletos, Whitney Newey, Harald Uhlig and Hal White.

on income expectations. We show that individuals' information sets can be characterized by simple nonparametric exclusion restrictions and that no assumption about agents' structural model is needed - in particular, we do not have to assume rational expectations.

In our application we find that individuals use rather coarse information to predict their future income. In particular, we are not able to reject that agents use only age and occupation status to predict future income growth. In contrast, their educational status, their sector of employment or their local labor market region is not contained in their predictions, once variation in age and occupational status is accounted for. After establishing which information individuals use when forming expectations, we test if these information sets are consistent with costless information processing, i.e. are agents able to productively use information as long as it is available to them. That agents might be constraint in the amount of information they can be attentive to has recently been claimed in the literature on rational inattention as pioneered by Christopher Sims (see e.g. Sims (2003); Mackowiak and Wiederholt (2000)). We first show that information processing costs cannot be identified without further restrictions on the structural model individuals use. We then devise a nonparametric test for the joint hypothesis of rational expectations and costless information processing.

While our method and the findings we obtain have implications for a wide range of situations, we study the consequences for individuals' consumption behavior. This is natural, as we analyze in particular individual's income expectations, and one of the main insights of the life-cycle model of consumption is that consumers' information about their earnings prospects crucially determines consumption expenditure through several channels. In particular, one of the central prediction of the life-cycle model is that individual smooth consumption intertemporally, and insure through other means as to keep utility constant. However, it is now widely accepted that in order to be quantitatively consistent with the empirical evidence, the model has to feature uninsurable income risk and risk aversion. When those features are added, there are incentives for precautionary savings. Once precautionary savings are important, however, consumers' information sets affect consumption through a mechanism which is distinct from the consumption smoothing mechanism: As better information about future income affects the perceived income variance, consumers' information sets also determine their incentives to engage in precautionary savings and hence consumption behavior.

To quantify the importance of individuals' information sets on consumption behavior, we use a standard life-cycle model with uninsurable labor income risk (Carroll, 1997; Deaton, 1991; Gourinchas and Parker, 2002). Through the lens of the model, consumers' information sets affect the agents' perceived environment in that they determine how much of the income process is predictable and how much has to be attributed to permanent and transitory shocks. Using the information sets as estimated from the microdata, we find that households overestimate the variance of transitory shocks compared with the econometrician and slightly underestimate the predictable rate of income growth. This misconception of the income process they face will change individual behavior. At the estimated parameters, the utility loss of excluding information from their information sets is small. While there there is heterogeneity in that poor individuals are harmed more, the average willingness to pay for the econometricians' information set amounts to roughly 0.05% of agents' permanent income. Hence, the information processing costs can be quite low for individuals to rationally chose to not incorporate different sources of information in their income predictions. The reason is that - in the model - occupational characteristics and age do a good job to decompose the observed time-series of income in the micro-data into predictable components and transitory and permanent shocks. With the individuals' model being close to the income process, the utility consequences are relatively small.

Related literature: Empirical studies of individuals' expectations in general and their information sets in particular, have a long tradition in economics. First of all, there is a large empirical literature that tests the rational expectations hypothesis (Lovell, 1986; Keane and Runkle, 1990; Brown and Matial, 1981). This literature has often tested for "informational efficiency", which is similar to our concept of costless information processing and hence closely related to our specification test. Secondly, there are numerous contributions that explicitly study subjective expectation data (Dominitz, 1998; Dominitz and Manski, 1997; Hurd and McGarry, 1995). While data on subjective expectations has often been met with skepticism, Manski (2004) provides evidence that such data is helpful to predict choices and argues that it should be used more often given its wide availability. Finally there is an extensive literature on forecasting, that models agents' forecasts as the solution of a well-defined maximization problem for given preferences and information sets (Pesaran and Weale, 2006; Machina and Granger, 2006).

Recently, expectations data have also been explicitly used for particular applications. Guiso et al. (1996) use agents' self-reported income uncertainty in a study of portfolio choice, Carroll (2003) exploits expectations on future inflation and unemployment rates to estimate a structural model of expectation formation and Jappelli and Pistaferri (2000) provide tests for consumption excess sensitivity when explicitly controlling for individuals' income expectations. Finally, Cunha et al. (2005) show how individual information sets can be recovered from a structural model of college choice in a life-cycle framework.

Regarding our application, the life-cycle of model of consumption is the workhorse model to analyze consumption behavior and has been tested extensively (see e.g. Hall and Mishkin (1982); Hall (1978); Attanasio and Weber (1995) and Browning and Lusardi (1996) for a review). While the robust finding that the observed consumption sensitivity to income shocks exceeds the one predicted by the standard model of perfect foresight (or its certainty-equivalent version with quadratic utility) and that changes in consumption are positively correlated with anticipated income shocks (the "excess sensitivity puzzle"), have often been interpreted as evidence against the life-cycle model, this conclusion has been challenged in the last decade. In particular, neither of these findings is inconsistent with the life-cycle theory once once uninsurable income uncertainty and risk-aversion is allowed for (Carroll, 2001, 1997). The importance of the precautionary savings motive to reconcile the empirical evidence with the life-cycle theory of consumption already suggests two crucial ingredients for individuals' consumption behavior. The first concern the income process itself, i.e. what are the statistical properties of the income process consumers face? In the context of the life-cycle model, many recent contributions use both consumption and income data simultaneously to learn about the structure of individual income (Gourinchas and Parker, 2002; Blundell et al., 2008; Krueger and Perri, 2011; Guvenen, 2007). The second one concerns consumers' information sets when forecasting future income. As the amount of information used when forecasting future income determines consumers' income uncertainty, the size of consumers' information sets might substantially affect consumption behavior if precautionary motives are important.[1] To use microdata on income expectations to learn about consumers' information sets when making consumption choices is the objective of this paper.

The structure of the paper is as follows. In the next section we will present our methodology to characterize information sets and give conditions for identification. Section three contains a tractable version of the life-cycle model to study the relation between consumers' information sets and consumption behavior and to characterize the demand for information. We then turn to our application. In section four we apply our econometric technique to microdata on income expectations and measure what information individuals use when forecasting future income. In section five we quantify the economic importance of agents' information on consumption behavior in the context of a standard life-cycle model. Section six concludes.

2 Characterizing Information Sets

To study these question we consider the following economy. There is a continuum of agents, which we model as realizations of a underlying random vector W. In particular, let $W = [X, U]$, where X is observable to the econometrician and U is unobservable. We are interested in agents' expectations about individual income Y, which is given by the structural relationship

$$Y = \psi(W) = \psi(X, U). \tag{1}$$

Individual (x, u) therefore earns income $y = \psi(x, u)$. Examples for $[X, U]$ are individual characteristics like education, experience or the match quality between the individual and the employer and aggregate characteristics like relative skill supplies in the individuals' local labor

[1]This is also the main difference to the paper of Luo (2008), who also studies consumption of inattentive consumers. However, he restricts the analysis to the CEQ version of the life-cycle model (i.e. he assumes quadratic utility) so there is no demand for precautionary savings.

market. However, U may also contain deeper concepts, like preferences, abilities etc., which we collectively refer to as "types". In sum, the objective distribution of income is fully determined by the structural relationship $\psi(.)$ and the underlying distribution of observables and types, $F_{X,U}$.

Individuals in this economy try to forecast future income. This need to forecast income arises because some realizations of $[X, U]$ might be unknown to the individual. This might for example be the quality of the future match if there is a chance that the current employment relationship might be terminated. The structural model agents have in mind is given by

$$Y = \psi^I(W) = \psi^I(X, U), \tag{2}$$

where ψ^I does not necessarily equal ψ. Note that it is without loss of generality to define both ψ and ψ^I over $[X, U]$ as we can always have either ψ or ψ^I to be trivial in the respective argument. It is in this sense that the function ψ^I could depend on individual specific variables: they could form part of U, and the objective function ψ could not depend on them.

To forecast their income, individuals employ "information" , which we model as a set of random variables Q. More precisely, the information is $\sigma(Q)$, where σ denotes the sigma-algebra spanned by Q, and we also denote the information set sometimes as $\mathcal{F}_Q$. However, for simplicity, we will mostly refer to Q as the information individuals hold.

Given this information, we denote the objective conditional distribution of $[X, U]$ given Q by $F_{X,U|Q}$. Analogously to above, the joint distribution of $[X, U]$ given Q as perceived by the agents is denoted $F^I_{X,U|Q}$; the same remark about heterogeneity applies. To emphasize, $F^I_{X,U|Q}$ and $F_{X,U|Q}$ do not have to be equal. We denote the respective densities by $f^I_{X,U|Q}$ and $f^I_{X,U|Q}$. Given Q and their "view of the world" encapsulated in $(F^I_{X,U|Q}, \psi^I)$, agents' subjective distribution of future income is given by

$$G^I_{Y|Q}(y, q) = \int_{(x,u):\psi^I(x,u)\leq y} f^I_{X,U|Q}(x, u; q)dxdu. \tag{3}$$

Hence, $G^I_{Y|Q}$ represents the joint mental act of having a structural relationship for Y given the characteristics (X, U) (i.e. ψ^I), and some process of learning about (X, U) from Q (i.e. $F^I_{X,U|Q}$). It is important to realize that, by observing realizations of $G^I_{Y|Q}(y, Q)$ which is a random variable in the cross-section, we cannot separately identify ψ^I and $F^I_{X,U|Q}$, at least not without further restrictions. Hence, we directly focus on $G^I_{Y|Q}$, and define this subjective distribution agents hold as agents' *forecasting model*, or *model.*

Definition 1. A model is a conditional distribution function of Y given information Q, i.e. a model is

$$G^I_{Y|Q}(y, q) = P[Y \leq y|Q = q; \psi^I, F^I] = \int_{-\infty}^{y} g^I_{Y|Q}(\eta; q)d\eta, \tag{4}$$

where g^I is derived from (3). In particular, G is a primitive of the individual's problem and is defined for all possible Q.

Similarly to (4), the econometricians beliefs about future income given the information Q (the econometrician's model) are given by

$$G_{Y|Q}(y,q) = P[Y \leq y|Q = q; \psi, F] = \int_{(x,u):\psi(x,u)\leq y} f_{X,U|Q}(x,u;q)dxdu. \tag{5}$$

Equation (4) and (5) illustrate that one may think of a forecasting model as a production function. It generates outputs (beliefs about future events) upon usage of inputs (information). This becomes apparent in the following example:

Example 1. Suppose that the structural relationship ψ is given by

$$Y = \psi(W) = a_X X + a_U U, \tag{6}$$

where X and U are scalars and individuals know the true model, i.e. $\psi^I = \psi$. Individuals' prior about X is given by $X \sim \mathcal{N}(\mu_X, \frac{1}{\rho_X})$ and they receive a signal $Q_X = X + \epsilon_X$, where $\epsilon_X \sim \mathcal{N}(0, \frac{1}{\eta_X})$. The information structure for U is analogous, and (X,U) and (ϵ_X, ϵ_U) are independent. Hence, individuals' beliefs about $[X,U]$ given Q are given

$$\begin{aligned} f^I_{X,U|Q}((x,u);(q_X,q_U)) &= f^I_{X|Q_X}(x;q_X)f^I_{U|Q_U}(u;q_U) \\ &= \frac{1}{\sigma_{X|Q_X}}\phi\left(\frac{x-\mu_{X|Q_X}(q_X)}{\sigma_{X|Q_X}}\right)\frac{1}{\sigma_{U|Q_U}}\phi\left(\frac{u-\mu_{U|Q_U}(q_U)}{\sigma_{U|Q_U}}\right), \end{aligned}$$

where ϕ is the normal density and $\mu_{X|Q_X}(q_X)$ and $\sigma^2_{X|Q_X}$ are the posterior mean and variance of X given Q_X, which are given by

$$\begin{aligned} \mu_{X|Q_X}(q_X) &= \frac{\rho_x\mu_X + \eta_x q_X}{\rho_x + \eta_x} \\ \sigma^2_{X|Q_X} &= \frac{1}{\rho_x + \eta_x}. \end{aligned}$$

$\mu_{U|Q_U}(q_U)$ and $\sigma^2_{U|Q_U}$ are defined analogously. The model G^I is then given by

$$\begin{aligned} G^I_{Y|Q}(y,(q_X,q_U)) &= P[Y \leq y|Q = q; \psi^I, F^I] \qquad (7) \\ &= \Phi\left(\frac{y - \sum_{i=X,U} a_i \frac{\rho_i\mu_i + \eta_i q_i}{\rho_i+\eta_i}}{\sqrt{\sum_{i=X,U}\frac{a_i^2}{\rho_i+\eta_i}}}\right) \qquad (8) \end{aligned}$$

Hence, the parameters of the agents' structural relationship ψ^I (i.e. $[a_i]_{i=X,U}$) and the parameters of the learning process (i.e. $[(\mu_i, \rho_i, \eta_i)]_{i=X,U}$) are not separably identified.[2]

Hence: once the individual has access to some information set $\sigma(Q)$, the beliefs for Y can be formed using $G^I_{Y|Q}$ costlessly. The population of individuals and their accompanying income expectations $G^I_{Y|Q}$ are therefore induced by realizations of the underlying random variable Q. From the point of view of the econometrician, it is precisely these observations of the random variable $G^I_{Y|Q}(Q; y)$, which we observe for all y. The two questions we ask are: (1) Are the income expectations we observe in the data consistent with a model of costly information processing[3]? (2) Can we tell theories of costly information processing apart from theories where individuals simply have a misspecified model of the world? To do so, we advance in two steps. First we simply ask what information individuals actually do use. Then we study in how far the chosen information sets are consistent with models of costly information processing.

2.1 What information do individuals use?

In our setup, the question of what information individuals use is formalized by asking which Q that span the information set $\sigma(Q)$ individuals *use*? The formal definition of what it means for information to be used is contained in the following definition.

Definition 2. Let $Q = [Q_1, Q_2]$. We say that the information in Q_2 is used conditional on Q_1, whenever for (y, q_1) with positive probability[4]

$$G^I_{Y|Q}(y; (q_1, q_2)) \neq G^I_{Y|Q}(y; (q_1, q_2'))$$

In words, information is used actively, whenever it affects the beliefs of some individuals in the population. We especially want to stress that usage of information is property of both the structural relationship for Y and the learning process that maps realizations of Q into beliefs about $[X, U]$. This can already be expected given the construction of $G^I_{Y|Q}$ in (3). It is also

[2]To see this, suppose that $a_U = 0$. Then (8) can be written as

$$G_{Y|Q}(y, q_X) \quad = \quad \Phi\left(\beta_0 y - \beta_1 - \beta_2 q_X\right),$$

where $\beta_0 = \frac{\sqrt{\rho_X + \eta_X}}{a_X}$, $\beta_1 = \frac{\rho_X \mu_X}{\sqrt{\rho_X + \eta_X}}$ and $\beta_2 = \frac{\eta_X}{\sqrt{\rho_X + \eta_X}}$. From the three β's we can not identify the four parameters $(a_X, \rho_X, \eta_X, \mu_X)$.

[3]We will define costly information processing formally below. Intuitively, we will think of information processing being costly if individuals face costs sampling Q.

[4]We say "(y, q_1) with positive probability", when all (y, q_1) form a set $\mathcal{Y}_1 \times \mathcal{Q}_{11} \subseteq \mathcal{Y} \times \mathcal{Q}_1$, with $P[\mathcal{Y}_1 \times \mathcal{Q}_{11}] > 0$, and analogously throughout this paper.

clearly seen in Example 1. According to Definition 2, we have that the information in Q_U is used conditional on Q_X, if for some y (see 8)

$$\Phi\left(\frac{y - a_X\frac{\rho_X\mu_X+\eta_X q_X}{\rho_X+\eta_X} - a_U\frac{\rho_U\mu_U+\eta_U q_U}{\rho_U+\eta_U}}{\sqrt{\sum_{i=X,U}\frac{a_i^2}{\rho_i+\eta_i}}}\right) \neq \Phi\left(\frac{y - a_X\frac{\rho_X\mu_X+\eta_X q_X}{\rho_X+\eta_X} - a_U\frac{\rho_U\mu_U+\eta_U q_U'}{\rho_U+\eta_U}}{\sqrt{\sum_{i=X,U}\frac{a_i^2}{\rho_i+\eta_i}}}\right),$$

with positive probability. This is the case if $\frac{a_U\eta_U}{\rho_U+\eta_U} \neq 0$. Hence, Q_U is *used* by individuals if *both* the information is considered informative ($\eta_U > 0$) and the factor it is predicting is part of the structural relationship ψ_I, i.e. $a_U \neq 0$. In contrast, Q_U is not used if either knowing Q_U does not help in predicting U (i.e. $\eta_U = 0$) or U is thought to be unrelated to income ($a_U = 0$).

When trying to characterize individuals information sets, we have to allow for the fact that individuals use information which is unobservable to the econometrician. Hence, we will only be able to make statements about variables, which are observable to us. Note that we can always write individual information sets as

$$Q = \pi(Z, V), Z \perp V$$

where Z is observed and V is unobserved, This construction of V being orthogonal to Z is exactly the right construction to characterize informational *content*. This does not mean that individuals actually use the observable variables Z but as long as there is a positive correlation between the observable variable Z and information they use, the information contained in Z is contained in the forecast. The following example makes this clear:

Example 2. Suppose income growth is a function of tenure T and ability A, i.e.

$$Y = \alpha T + \gamma A + \eta, \tag{9}$$

where η is a mean zero error term, tenure is observed by the econometrician (i.e. $T = X$) and ability is unobserved ($[A, \eta] = U$). Individuals base their forecast on (T, A), i.e. $Q = [T, A]$. Now suppose we were to ask if individuals use information about tenure (T) and educational attainment (E) when forecasting income. We can always write $A = h(T, E, V)$, where $V \perp T, E$, so that $Q = [T, A] = [T, h(T, E, V)] = \pi(Z, V)$, where $Z = [T, E]$. If ability and education are correlated conditional on tenure, then $h(T, E, V)$ is non-trivial in E and we will correctly find that the information in education is reflected in individual forecasts. We as econometricians cannot say if there is a Mincerian skill premium or if such skill premium is purely spurious (in the example (9) obviously the latter is the case). But to measure informational content we are not interested in the underlying structural model. For us it is only important if the information contained Z (here education) is reflected in individuals' forecasts. This example also makes clear why Definition 2 makes only conditional statements. If we were to observe ability, then we would correctly conclude that educational information is not used conditional on ability and tenure.

We now analyze the individual predictions more formally. To this end, we assume that (Y, Z, V) are jointly continuously distributed. Individuals' beliefs about their future income can therefore be written as nonseparable model, i.e.,

$$G^I_{Y|Q}(y;q) = G^I_{Y|\pi(Z,V)}(y;z,v) = \int_{(x,u):\psi^I(x,u)\leq y} f^I_{X,U|\pi(Z,V)}(x,u;z,v)dxdu \tag{10}$$

$$\equiv \varphi(z,v;y). \tag{11}$$

We emphasize here that we think of y as a fixed index, and of Z and V as the actual argument of the function, i.e., $\varphi(Z,V;y)$ denotes the (random) conditional probability that $Y < y$, varying across the population as the individuals information $Q = (Z, V)$, varies across the population. Since we do not care about the structural relationship, we can choose the unobservable V to be independent of Z, and for any y, z, enter $\varphi(z, \cdot; y)$ strictly monotonically. From Matzkin (2003), w.l.o.g., we can let $V \backsim U[0,1]$, and can then conclude that there is a family of quantile regressions indexed by y, i.e.,

$$\varphi(z,v;y) = k^v_{G^I_y|Z}(z) \tag{12}$$

where $k^v_{G^I_y|Z}(z)$ denotes the v quantile of $G^I_{Y|Q}(y;Q)$.

In words, we can consider the cross-sectional quantiles of the individual predictions of probabilities for any value of interest y as a tool to evaluate what information they use. The v quantile of $G^I_{Y|Q}(y;Q)$ given $Z = z$ gives then the prediction of an individual with observable information $Z = z$, and unobservable reduced form information $V = v$. As people with different realizations of V given Z make different predictions for their income growth, $k^V_{G^I_y|Z}(z)$ is a random variable (given z) and the unobserved heterogeneity in the population is encapsulated in the random variable V. Basically: When we are not interested in the structural model, we can fully control for the unobserved heterogeneity by equating it to the quantiles of the reduced form distribution.

Example 3. To see why this construction works, consider Example 1 above, suppose for simplicity $\mu_X = \mu_U = 0$ and suppose that the individual observes two signals, where $Q_X = Z$ and $Q_U = V$, i.e. only the first information is observed by the econometrician. Then (8) implies that

$$G^I_{Y|Z,V}(y;z,v) = \Phi\left(\frac{y - \theta_z z - \theta_v v}{\delta}\right)$$

where $\delta = \sqrt{\sum_{i=X,U} \frac{a_i^2}{\rho_i+\eta_i}}$ and $\theta_i = \frac{a_i\eta_i}{\rho_i+\eta_i}$. Now define $M = \Phi(\frac{V}{\sigma_V})$ so that $M \sim U[0,1]$ and

$$G^I_{Y|Z,M}(y;z,m) = \Phi\left(\frac{y - \theta_z z - \theta_v \sigma_V \Phi^{-1}(m)}{\delta}\right).$$

The α quantile of $G^I_{Y|Q}$ given $Z = z$, $k^\alpha_{G^I_y|Z}(z)$, is then given by

$$\begin{aligned}
\alpha &= P\left[\Phi\left(\frac{y - \theta_z z - \theta_v \sigma_V \Phi^{-1}(M)}{\delta}\right) \leq k^\alpha_{G^I_y|Z}(z)|Z = z\right] \\
&= 1 - P\left[M \leq \Phi\left(\frac{y - \theta_z z - \delta\Phi^{-1}(k^\alpha_{G^I_y|Z}(z))}{\sigma_V\theta_v}\right)|Z = z\right] \\
&= 1 - \Phi\left(\frac{y - \theta_z z - \delta\Phi^{-1}(k^\alpha_{G^I_y|Z}(z))}{\sigma_V\theta_v}\right).
\end{aligned}$$

Hence, solving for $k^\alpha_{G^I_y|Z}(z)$ yields

$$\begin{aligned}
k^\alpha_{G^I_y|Z}(z)) &= \Phi\left(\frac{y - \theta_z z - \theta_v \sigma_v \Phi^{-1}(1-\alpha)}{\delta}\right) \\
&= G_{Y|Z,V}(y;z,\sigma_v\Phi^{-1}(1-\alpha)) \\
&= G_{Y|Z,M}(y;z,1-\alpha) \\
&\equiv \varphi(z,\alpha;y),
\end{aligned} \tag{13}$$

which is exactly the form of (12). Hence, for any v there is a specific α such that the conditional distribution of $G^I_{Y|Z,V}$ given $[Z,V]$ is equal to the α-quantile of the conditional distribution of G^I_y given Z.

This construction suggests how we can test for the informational content of individuals' forecasts, i.e. to answer the question if there is a positive measure of people paying attention to some information. In particular, let $Z = [Z_1, Z_2]$. Then, individuals do not use the information contained in Z_2 conditional on $[Z_1, V]$, if for all (y, z_1, v) we have

$$k^v_{G^I_y|Z_1,Z_2}(z_1,z_2) = k^v_{G^I_y|Z_1}(z_1) \tag{14}$$

If this was the case, individuals receiving the signal $q = (z_1, z_2, v)$ report the same income expectation as individuals receiving the signal $q = (z_1, v)$, i.e. individuals do not incorporate information contained in Z_2 once $[Z_1, V]$ is controlled for. As (14) contains our first testable restriction, we state this is in the following proposition.

Proposition 1. *Consider the model above. Let individuals' information sets be given by* $Q = \pi(Z, V)$. *Let* $Z = [Z_1, Z_2]$. *Then, individuals use* Z_2 *conditional on* $[Z_1, V]$ *in the sense of Definition 2 if and only if*

$$k^v_{G^I_y|Z_1,Z_2}(z_1, z_2) \neq k^v_{G^I_y|Z_1}(z_1)$$

for z_1 *with positive probability.*

Proof. Follows directly from Definition 2, (11) and (12). □

To see why (14) is the correct test for the usage of information, consider again the Example.

Example 4. Recall from (13) that

$$k^v_{G^I_y|Z_1,Z_2}(z_1, z_2) = \Phi\left(\frac{y - \theta_{z_1} z_1 - \theta_{z_2} z_2 - \theta_v \sigma_v \Phi^{-1}(1-v)}{\delta}\right).$$

$k^v_{G^I_y|Z_1}(z_1)$ however, is implicitly defined by

$$1 - v = E\left[\Phi\left(\frac{y - \theta_{z_1} z_1 - \theta_{z_2} Z_2 - \delta\Phi^{-1}(k^v_{G_y|Z_1}(z_1))}{\sigma_V \theta_v}\right) | Z_1 = z_1\right]$$

For these to be equal for all (y, v, z_1), we need that $\theta_{z_2} = \frac{a_{z_2}\eta_{z_2}}{\rho_2 + \eta_{z_2}} = 0$, which is exactly the condition that Z_2 is not used by individuals.

Hence, using the quantile exclusion restriction contained in Proposition 1, we can exactly characterize which information is affects individuals' beliefs.[5]

[5] While the quantile function $k^v_{G^I_y|Z}$ is exactly the right statistic to test for informational content, we can also look at the conditional mean function. Doing so delivers an intuitive but weaker test for informational usage. In particular, given agents' information Q and view of the world $(\psi^I, f^I_{X,U|Q})$, their future expected income is given by

$$\begin{aligned} E^I[Y|Q=q] = \int y g^I_{Y|Q}(y,q)dy &= \int \psi^I(w) f^I_{W|Q}(w;q)dw = \int \psi^I(w) f^I_{W|\pi(Z,V)}(w;(z,v))dw \\ &\equiv m(z_1, z_2, v), \end{aligned} \tag{15}$$

where again $Z = [Z_1, Z_2]$. If Z_2 is not used conditional on $[Z_1, V]$, then $m(z_1, z_2, v) = m(z_1, v)$, i.e. $m(z_1, z_2, v)$ is trivial in z_2, where m is defined in (15). Note that this is only an "if" statement. However, it is an "if and only if" statement under the regularity condition that changes in the subjective density $g^I_{Y|Q}(y, q)$ do not average out once we integrate over y. Hence, Proposition 1 is stronger because if focuses on this subjective distribution directly. Looking at the exclusion restriction contained in (15) is still useful because it can be tested on data, where only point estimates of agents' income expectations are available. Note also that (15) is can be constructed as a weighted average of the different quantiles $k^v_{G^I_y|Z}$, so that (15) contains strictly less information.

2.2 Do individuals use all valuable information or is there evidence for costly information processing?

Having characterized the content of individual information sets above, we now ask in what sense the finding that some variable is not part of individual information sets is evidence for costly information processing. In this setup, this can be rephrased as saying: Would someone endowed with the model $G^I_{Y|Q}$ but no information processing costs have chosen to use this information? Hence, the essence of costly information processing is that there is a demand for information, but that the marginal value falls short of the marginal processing costs. To test for costly information processing, we therefore have to define the value of (or demand for) information.

Definition 3. Consider the setup described above. Let $Q = [Q_1, Q_2]$. We say that the information contained in Q_2 is valuable given the model G^I and the information Q, whenever

$$G^I_{Y|Q}(y, (q_1, q_2)) \neq G^I_{Y|Q}(y, (q_1, q_2')) \tag{16}$$

with positive probability. For notational simplicity we will say that Q_2 is $(G^I, [Q_1, Q_2]) -$ *valuable* if (16) holds true.

Hence, according to Definition 3, additional information is valuable whenever it changes the individuals' posterior in some states of the world. While we think this definition being natural in our setup, we also want to stress that in general the demand for information obviously also depends on the preferences of the individual. If no decision depends on the beliefs about personal income (and the decision maker does not experience any utility loss from ambiguity aversion or other behavioral aspects), the demand for information is obviously zero as the individual does not care about her posterior beliefs about income. In Definition 3 we do not consider these possibilities, i.e. we only care about cases, where the individual actually cares, about the beliefs she ends up with, before decisions have to be taken. In the Appendix we show what properties the agent's utility function has to satisfy only some weak restriction for this definition of the value of information to be sensible.

Given this definition of information being valuable and our definition of information usage, we can also give a precise definition of what we are looking for in order to find costly information processing.

Definition 4. Consider the setup described above. We say that individuals are characterized by costly information processing with respect to Q_2, whenever Q_2 is not used conditional on Q_1, despite Q_2 being $(G^I, [Q_1, Q_2])$-valuable.

Hence, whenever some information Q_2 would have changed individuals' posterior (given their model and their information) but individuals decide to not use Q_2, we will conclude that their expectation formation process is subject to costly information processing. The important aspect of Definition 4 is precisely the dependence of the value of information on G^I and on $\sigma(Q_1)$ - both of which are unobserved by the econometrician. Therefore the question is: Can we detect occurrences of costly information processing given data on income expectations without further restrictions on G^I and $\sigma(Q_1)$?

As before let $Q = \pi(Z, V) = \pi(Z_1, Z_2, V)$, where Z_2 is not used conditional on $[Z_1, V]$. From Section 2.1 above, we can find those Z_2 in the data. Our first result is an impossibility result to find evidence in favor of processing costs without further restrictions on the agents' model.

Proposition 2. *Consider the setup above. Without further assumptions on G^I it is impossible to conclude that individuals are characterized by information processing costs.*

Proof. We can always find $(\psi^I, f^I_{X,U|Q})$ such that Z_2 is not $(G^I, \sigma(Z_1, Z_2, V))$-valuable. □

Proposition 2 is very intuitive: If the model agents are using is such that Z_2 is considered noise, Z_2 would not have been used even without processing costs. Hence, in order to give the hypothesis of costly information processing empirical content, we impose the following restriction on the relationship between the agents' and the objective model.

Assumption 1. *For all Q_1, Q_2, if Q_2 is $(G, [Q_1, Q_2])$-valuable, then Q_2 is also $(G^I, [Q_1, Q_2])$-valuable.*

Assumption 1 requires a minimum amount of consistency between the agents' view of the world and the structural model of the economy. Hence, we refer to Assumption 1 as an assumption of "weak rationality". Intuitively, it requires the following: whenever the econometrician with $Q = [Q_1, Q_2]$ at his disposal would not discard Q_2, we require that individuals would not do so either. Individuals could disagree with the econometrician how Q_2 enters, but they have to realize that Q_2 determines the distribution of income conditional on Q_1. We consider Assumption 1 to be very weak and it turns out that it is sufficient to detect costly information processing in the data.

To do so, note that observing realizations of $G_{Y|Q}(y; Q)$ for different y is the most we can learn about the information individuals have. Hence, one question we can ask is: Is Z_2, which recall is not used by individuals, valuable conditional on the individuals' perceived distribution of income? In particular, let $G^I_{Y|Q}(y; Q) = h_y(Q)$, which is indexed by y. Anticipating our application, let y take K finite values so that we observe realizations of $\{h_{y_j}(Q)\}_{j=1}^K$.[6]. As Y, Z_2

[6]At the expense of further notation we could generalize to consider y being continuous.

and $\{h_{y_j}(Q)\}_{j=1}^K$ is observed, the conditional distribution of Y, conditional on Z_2, $\{h_{y_j}(Q)\}_{j=1}^K$ is also observed, i.e. we as econometricians can form, for all $y \in \mathcal{Y}$,

$$P[Y \leq y | \{h_{y_j}(Q) = h_{y_j}(q)\}_j, Z_2 = z_2] \equiv \gamma(y, \{h_{y_j}(q)\}_j, z_2). \tag{17}$$

Then we get the following proposition.

Proposition 3. *Consider the setup described above and let γ be defined in* (17)*. Let $Z_2 \in \mathcal{F}$. Then $G^I = G$ if and only if*

$$\gamma(y, \{h_{y_j}(q)\}_j, z_2) = h_y(q) \quad \forall (y, z_2, z_2'),$$

i.e. γ is trivial in z_2. Let $Z_2 \neq \mathcal{F}$ and let Assumption 1 hold true. Then information processing is costly whenever

$$\gamma(y, \{h_{y_j}(q)\}_j, z_2) \neq \gamma(y, \{h_{y_j}(q)\}_j, z_2')$$

with positive probability.

Proof. Consider the first part. Note first that

$$\begin{aligned} P[Y \leq y | \{h_{y_j}(Q) = h_{y_j}(q)\}_j, Z_2 = z_2] &= E[P[Y \leq y | Q, Z_2] | \{h_{y_j}(Q) = h_{y_j}(q)\}_j, Z_2 = z_2] \\ &= E[G_{Y|Q,Z_2}(y; Q, Z_2) | \{h_{y_j}(Q) = h_{y_j}(q)\}_j, Z_2 = z_2] \end{aligned} \tag{18}$$

Again we express agents' information set as

$$Q = \pi(\{h_{y_j}(Q)\}_j, Z_2, S) \quad \text{with } S \perp \{h_{y_j}(Q)\}_j, Z_2. \tag{19}$$

Then

$$\begin{aligned} & E[G_{Y|Q,Z_2}(y; Q, Z_2) | \{h_{y_j}(Q) = h_{y_j}(q)\}_j, Z_2 = z_2] \\ = & E[G_{Y|\{h_{y_j}(Q)\}_j, Z_2, S}(y; \{h_{y_j}(Q)\}_j, Z_2, S) | \{h_{y_j}(Q) = h_{y_j}(q)\}_j, Z_2 = z_2] \\ = & \int G_{Y|\{h_{y_j}(Q)\}_j, Z_2, S}(y; \{h_{y_j}(q)\}_j, z_2, s) f_{S|\{h_{y_j}(Q)\}_j, Z_2}(s; \{h_{y_j}(q)\}_j, z_2) ds \\ = & \int G_{Y|\{h_{y_j}(Q)\}_j, Z_2, S}(y; \{h_{y_j}(q)\}_j, z_2, s) f_S(s) ds. \end{aligned} \tag{20}$$

If $G^I = G$, then

$$\begin{aligned} & E\left[1\{Y < y_s\} | Z_2, G_{Y_j|Q}, j = 1, ..., K\right] \\ = & E\left[E\left[1\{Y < y_s\} | Q, Z_2, G_{Y_j|Q}, j = 1, ..., K\right] | Z_2, G_{Y_j|Q}, j = 1, ..., K\right] \\ = & E\left[E\left[1\{Y < y_s\} | Q\right] | Z_2, G_{Y_j|Q}, j = 1, ..., K\right] \\ = & G_{Y_s|Q}. \end{aligned}$$

Similarly, if (20) is equal to $h_y(q)$, then

$$G_{Y|\{h_{y_j}(Q)\}_j,Z_2,S}(y;\{h_{y_j}(q)\}_j,z_2,s) = h_y(q), \tag{21}$$

which generically is only true if $G^I = G$. Now consider the second part. If γ is not trivial in z_2, then (20) implies that

$$G_{Y|\{h_{y_j}(Q)\}_j,Z_2,S}(y;\{h_{y_j}(q)\}_j,z_2,s) \neq G_{Y|\{h_{y_j}(Q)\}_j,Z_2,S}(y;\{h_{y_j}(q)\}_j,z_2',s) \tag{22}$$

for some (z_2, z_2') with positive probability. Under (19) this implies that Z_2 is $(G,[Q,Z_2])$-valueable. Under Assumption 1 this also implies that Z_2 is $(G^I,[Q,Z_2])$-valuable. As $Z_2 \notin \mathcal{F}$, information processing is costly. □

In essence, Proposition 3 shows the following. Whenever we as econometricians would consider some information Z_2 valuable conditional on knowing agents' forecasts, the assumption of weak rationality implies that agents would consider that information valuable as well. This is inconsistent with costless information processing. Similarly, whenever we know that agents do use Z_2 in their forecast, this information should only have predictive power on the econometrician's forecast whenever the agents' model differs from the one of the econometrician.

To illustrate the content of Proposition 3 consider again our example.

Example 5. Suppose the true structural relationship is given by $Y = \psi(X,U) = a_X X + U$. In contrast, the relationship individuals perceive is $Y = \psi^I(X,U) = a_X^I X + U$. The objective distribution of (X,U) is as above, i.e. they are independently normally distributed with means (μ_X,μ_U) and precisions (ρ_X,ρ_U). While individuals are assumed to know ρ_i, their perceived prior means are (μ_X^I,μ_U^I). Individuals receive two signals Q and Z, which are given by $Q = X + \epsilon_Q$ and $Z = Q + \lambda X + \epsilon_Z$. Individuals do not have any information about U. The error terms (ϵ_Q,ϵ_Z) are independently normally distributed, have mean zero and precisions (η_Q,η_Z). Note that λ parametrizes the informational value of Z conditional on Q. If $\lambda = 0$, Z does not contain any valuable information about X conditional on observing Q. Observing Z and Q is clearly equivalent to observing Q and $\hat{Z} = \frac{Z-Q}{\lambda}$, which has mean X and precision $\lambda\eta_Z$.[7] The econometrician's model (conditional on observing $(Q,\hat{Z})$) is given by

$$G_{Y|Q,\hat{Z}}(y;Q,\hat{Z}) = \Phi\left(\frac{y-\theta_\mu-\theta_Q Q-\theta_{\hat{Z}}\hat{Z}}{\delta}\right), \tag{23}$$

[7] $\tilde{Z}$ is of course only defined for $\lambda \neq 0$. We will be taking the limit later on.

where $\delta = \sqrt{\frac{1}{\rho_U} + \frac{a_X^2}{\rho_X + \eta_Q + \lambda\eta_Z}}$, $\theta_\mu = \mu_U + a_X \frac{\rho_X \mu_X}{\rho_X + \eta_Q + \lambda\eta_Z}$, $\theta_Q = a_X \frac{\eta_Q}{\rho_X + \eta_Q + \lambda\eta_Z}$ and $\theta_{\hat{Z}} = a_X \frac{\lambda\eta_Z}{\rho_X + \eta_Q + \lambda\eta_Z}$. The agents' model, who use only Q when forecasting their income, is given by

$$G^I_{Y|Q}(y;Q) = \Phi\left(\frac{y - \theta^I_\mu - \theta^I_Q Q}{\delta^I}\right), \tag{24}$$

where $\delta^I = \sqrt{\frac{1}{\rho_U} + \frac{(a^I_X)^2}{\rho_X + \eta_Q}}$, $\theta^I_\mu = \mu^I_U + a^I_X \frac{\rho_X \mu^I_X}{\rho_X + \eta_Q}$ and $\theta^I_Q = a^I_X \frac{\eta_Q}{\rho_X + \eta_Q}$.

Now consider γ defined in (17). For this example we have

$$\begin{aligned}
\gamma(h_y(q), z) &= E[G_{Y|Q,Z}(y;Q,Z)|h_y(Q) = h_y(q), Z = z] \\
&= E\left[\Phi\left(\frac{y - \theta_\mu - \theta_Q Q - \theta_{\hat{Z}}\hat{Z}}{\delta}\right)|\Phi\left(\frac{y - \theta^I_\mu - \theta^I_Q Q}{\delta^I}\right) = h_y(q), Z = z\right].
\end{aligned} \tag{25}$$

If consumers have rational expectations, $a^I_X = a_X$, $\mu^I_X = \mu_X$ and $\mu^I_U = \mu_U$. If furthermore consumers do not face any information processing costs, Z cannot be (G,Q)-valuable as Z is not used. Definition 3 therefore implies that $G_{Y|Q}(y;Q) = G^I_{Y|Q,\hat{Z}}(y;Q,\hat{Z})$ with probability one, so that

$$\Phi\left(\frac{y - \tilde{\theta}_\mu - \tilde{\theta}_Q q}{\tilde{\delta}}\right) = \Phi\left(\frac{y - \theta_\mu - \theta_Q q - \theta_{\hat{Z}}\hat{z}}{\delta}\right) \quad \forall(y,q,z), \tag{26}$$

where $\tilde{\delta} = \sqrt{\frac{1}{\rho_U} + \frac{a_X^2}{\rho_X + \eta_Q}}$, $\tilde{\theta}_\mu = \mu_U + a_X \frac{\rho_X \mu_X}{\rho_X + \eta_Q}$ and $\tilde{\theta}_Q = a_X \frac{\eta_Q}{\rho_X + \eta_Q}$. For (26) to hold true, we need that $\theta_{\hat{Z}} = 0$, $\tilde{\delta} = \delta$, $\tilde{\theta}_Q = \theta_Q$ and $\tilde{\theta}_\mu = \theta_\mu$. This is the case if and only if $\lambda = 0$, i.e. Z does not contain any information about X conditional on Q. But then, (25) implies that

$$\begin{aligned}
\gamma(h_y(q), z) &= E\left[\Phi\left(\frac{y - \tilde{\theta}_\mu - \tilde{\theta}_Q Q}{\tilde{\delta}}\right)|\Phi\left(\frac{y - \tilde{\theta}_\mu - \tilde{\theta}_Q Q}{\tilde{\delta}}\right) = h_y(q), Z = z\right] \\
&= E\left[h_y(q), Z = z\right] = \gamma(h_y(q)),
\end{aligned}$$

i.e. γ is trivial in z. It on the other hand, agents have rational expectations, but information processing was not costless, it is possible that agents exclude Z despite $\lambda \neq 0$, so that

$$\gamma(h_y(q), z) = E\left[\Phi\left(\frac{y - \theta_\mu - \theta_Q Q - \theta_{\hat{Z}}\hat{Z}}{\delta}\right)|\Phi\left(\frac{y - \tilde{\theta}_\mu - \tilde{\theta}_Q Q}{\tilde{\delta}}\right) = h_y(q), \hat{Z} = \hat{z}\right],$$

which is not trivial in $\hat{z}$. Similarly, suppose that agents have a misspecified model of the world and consider Z not valuable given Q. In particular, suppose that agents perceive that $\eta_Z^I = 0 < \eta_Z$ and that $a_X^I \neq a_x$. Then,

$$\gamma(h_y(q), z) \quad = \quad E\left[\Phi\left(\frac{y - \theta_\mu - \theta_Q Q - \theta_{\hat{Z}} \hat{Z}}{\delta}\right) |\Phi\left(\frac{y - \theta_\mu^I - \theta_Q^I Q}{\delta^I}\right) = h_y(q), \hat{Z} = \hat{z}\right],$$

which is generically a function of $\hat{z}$. The same holds of course true if individuals have neither rational expectations, not costless information processing.

3 Empirical Analysis

In this section we will apply this framework to cross-sectional micro-data on individuals' income expectations. As in the theory laid out in Section 2, we will first measure the content of individual information sets and then ask if the micro-data is consistent with models of costly information processing.

3.1 Data Sources

The data we use is from the 'Survey of Household Income and Wealth' (SHIW), collected by the Bank of Italy.[8] The SHIW provides detailed information on individual characteristics, sources of income, and financial assets for about 8000 households (roughly 24.000 individuals). In 1991, the survey included a question on individual income and inflation expectations. The same data was also used in Jappelli and Pistaferri (2000), who use the expectation data as an instrument for consumption growth in a standard Euler equation framework and in Guiso et al. (1996), who show that income expectations are helpful in explaining portfolio choices. The question about individual income expectations has the following wording: "Think about your entire working income or pension payments. On this card you see several possible categories of growth rates. Which possibilities concerning your income change do you rule out? Assume you could distribute 100 points on the remaining categories: how many points would you give to each category?" The 12 categories spanned an interval between 0% and 25% with the intervals in the center being tighter than on the boundaries. We construct individual income forecasts from this data by forming

$$Y^*(\mathcal{F}_i) = \left(\sum_{c=1}^{12} p_c^i \mu_c\right) I_i, \tag{27}$$

[8]The data and all the programs used to generate the results of this paper are available on our website.

where μ_c is the midpoint of the interval c, p_c^i denotes the probability of income growth being in interval c as stated by individual i and I_i denotes current income of individual i. We conducted the analysis only for the 1991 sample as there are more observations with valid expectation data.

Besides the expectation data, the SHIW survey also contains data on various economic characteristics. It will be those characteristics for whose exclusion we will test. Note especially the entire data is self-reported, i.e. our analysis does not suffer from the problem that individuals might not have access to the information the researcher tests for. So if we conclude that some variable Z is not included in the income expectations, we can rule out the case that Z was not known to the individuals. They clearly knew Z but decided to not use it when forming their income expectations. This aspect of the data is important because it allows us to exclusively focus on the aspect the literature on rational inattention focuses on - in principle individuals have access to a wide range of information but they optimally chose to be inattentive to parts of it.

3.2 Descriptive Statistics and Reduced Form Results

Before turning to the nonparametric test, we take a reduced form look at the data to gauge the validity of the reported income expectations. In Table 1 below we regress the realized income growth (for both labor and capital income) on individuals' expectations and other characteristics. We see that there is a robust positive correlation between expected and realized income growth for labor income. In its purest form, Table 1 could be considered as a test for the rational expectations hypothesis, according to which we would expect a coefficient of unity on individuals' expected income growth and a coefficient of zero on other characteristics. However, as stressed by Keane and Runkle (1990), this requires the assumption that there are no aggregate shocks, which is unlikely to be the case for our application. In any case, we are *not* testing for the rationality of individuals' expectations but are only concerned with the size of their information sets. Hence, we view Table 1 as reassuring that individuals' reported income expectations are not merely noise but in fact do have predictive power for realized growth rates. Additionally, the results also show that individuals seem to predict their labor income and not their capital income. In the last two columns we regress the growth rate of capital income on individuals' expectations and do not find any discernible pattern, because the coefficients are very imprecisely estimated. As many individuals do not report any capital income, the last column focuses only on individuals reporting non-zero capital income growth. The standard errors decline substantially and the estimated coefficient is statistically zero.

Now consider a first pass to measure the informational content of individual information sets. In Table 3.2 below we report the results of simple OLS regression of individuals' expected income growth, i.e. $\int \gamma g_{\Gamma|Q}(\gamma; q)\, d\eta$, where Q denotes the information sets and $g_{\Gamma|Q}$ denotes the agents'

			Realized Growth of Laor Income			Realized Growth of Capital Income	
exp_growth	0.517**	0.637**	0.504**	0.336**	0.112**	4.080	-0.224
	(0.114)	(0.170)	(0.120)	(0.101)	(0.0491)	(4.197)	(0.353)
lnwages		-0.233*	-0.382**	-0.482**	-0.145**	-0.187	-0.0600
		(0.139)	(0.194)	(0.226)	(0.0113)	(0.387)	(0.0572)
age		0.0195*	0.0266**	0.00594	0.0112**	0.181**	0.00974
		(0.0110)	(0.00706)	(0.00832)	(0.00221)	(0.0919)	(0.0123)
agesq		-0.000259**	-0.000329**	-0.0000389	-0.000122**	-0.00230**	-0.000106
		(0.000127)	(0.0000867)	(0.000106)	(0.0000260)	(0.00109)	(0.000137)
N	2075	2075	2075	1827	1665	1841	1571
R^2	0.004	0.066	0.115	0.168	0.132	0.005	0.015

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$

Notes: Robust standard errors in parentheses. Specifications 3 to 7 control for a full set of education, industry, occupation and area fixed effects.

Table 1: Predictive Power of Income Expectations

conditional distribution of income growth rates, on different cross-sectional characteristics. If the true model was indeed linear, this specification would indeed be a consistent test for the size of Q. Table 3.2 provides us with a reduced form sense, which information individuals do pay attentions and which not. While current (log) income, local labor market conditions (which we think of being captured by the area dummies) and occupational characteristics are highly significant and therefore not excluded from individuals' income expectations, age and education are not part of individual information sets. In the following we will test these hypothesis non-parametrically as required by the theory.

3.3 Testing for Informational Content

We now turn to the consistent test of individuals' information sets. According to Proposition 1, we can test for the inclusion of Z_2 conditional on Z_1 by testing the nonparametric exclusion restriction

$$k^v_{G^I_y|Z_1,Z_2}(z_1, z_2) \neq k^v_{G^I_y|Z_1}(z_1) \quad \text{for all } z_2, \tag{28}$$

where $k^v_{G^I_y|Z_1,Z_2}(z_1, z_2)$ is the vth quantile of G^I. At this point, we implemented the test only for the mean regression, i.e. we test for the exclusion of Z_2 in the conditional mean function

$$m_{E^I[Y]}(z_1, z_2) = m_{E^I[Y]}(z_1) \quad \text{for all } z_2, \tag{29}$$

where $m_{E^I[Y]}(z_1, z_2)$ denotes the conditional mean of individuals' expected income. To estimate the restriction imposed in (29), we take the following procedure (the details are contained in the Appendix):

1. We estimate first estimate $m_{E^I[Y]}(z_1)$ using a standard local linear non=parametric estimator

2. Given the estimates $\hat{m}_{E^I[Y]}(.)$, we construct the the residuals $\hat{\varepsilon}_i = E^I[Y_i] - \hat{m}_{E^I[Y]}(z_{1,i})$

3. With these residuals at hand, we construct B bootstrap samples $E^I[Y_i]^* = \hat{m}_{E^I[Y]}(z_{1,i}) + \hat{\varepsilon}^*_i$, where $\hat{\varepsilon}^*_i$ are bootstrap residuals, which we obtain from $\hat{\varepsilon}_i$ using the wild bootstrap proposed in Haerdle and Mammen (1993). Note that $E^I[Y_i]^*$ is constructed *under the null* that Z_2 is in fact excluded from individuals' information sets.

4. We then calculate the teststatistic

$$\tau = \int [m_{E^I[Y]}(z_1) - m_{E^I[Y]}(z_1, z_2)]^2 \omega(z_1), \tag{30}$$

	(1)	(2)	(3)	(4)	(5)	(6)
	exp_growth	exp_growth	exp_growth	exp_growth	exp_growth	exp_growth
lnwages	0.00611**	0.00342	0.00112	0.00777**	0.00657**	0.00270
	(0.00243)	(0.00281)	(0.00295)	(0.00253)	(0.00255)	(0.00317)
age	0.000332	0.000420	0.000452	0.000201	0.000348	0.000350
	(0.000604)	(0.000606)	(0.000602)	(0.000604)	(0.000603)	(0.000608)
agesq	-0.00000796	-0.00000860	-0.00000945	-0.00000688	-0.00000811	-0.00000841
	(0.00000706)	(0.00000706)	(0.00000705)	(0.00000705)	(0.00000706)	(0.00000708)
	1	2	3	4	5	6
F-Test: Education		1.66 (0.155)				0.46 (0.76)
F-Test: Occupation			6.34 (0.00)			4.16 (0.01)
F-Test: Region				7.56 (0.00)		7.16 (0.00)
F-Test: Sector					1.74 (0.156)	3.23 (0.0214)
N	3196	3196	3196	3196	3196	3196
R^2	0.008	0.010	0.014	0.017	0.009	0.025

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$

Table 2: Individual Information Sets: Reduced Form Estimates

where $\omega(z_1)$ is a weighting function and we estimate τ by its sample counterpart.[9]

5. Using the B bootstrap samples we then use (30) to estimate the distribution of τ, say H_τ
6. We then conclude that Z_2 is excluded from individuals' information sets conditional on Z_1 if τ does not exceed the 95% quantile of H_τ.

The results of this exercise are contained in Table 3 below, where each row contains the test results for one model. Consider for example the first row. There we considered the null hypothesis

$$\mathcal{F}_0 = \sigma\left(ln\left(y\right), age, education\right),$$

against the alternative

$$\mathcal{F}_A^1 \quad = \quad \sigma\left(\mathcal{F}_0, sector, area, occupation\right).$$

Hence, the first row contains the results of a test if sectoral, regional *and* occupational characteristics can be excluded from individuals' information sets once age, current income and educational characteristics are controlled for. The third column contains the teststatistic (calculated according to (30)), the fourth column the critical value, i.e. the 95% quantile of the distribution of the teststatistic. The last column finally contains the p-value. The first row therefore shows we can confidently reject the hypothesis that neither occupations, regional or sectoral characteristics matter for individuals' predictions, once income, age and education is controlled for. Row two shows that this rejection is not only driven by firms' occupational characteristics, but that either regional labor markets or sectoral affiliation matter independently. Row three finally shows that individuals do not seem to include their sector of employment in their information set. According to these results, agents include their occupation in their information set. Rows 4 to 6 therefore include the occupational status in agents' information set and show that now neither regional nor sectoral information seems to be important. Rows 7 to 9 now test for the direct impact of educational characteristics once occupational status is controlled for and show that individuals' information set are well described by only containing income, age and occupational status. These results are also visible in Figure 1 , where we depict the distribution of the teststatistic, the critical value (dashed line) and the value of teststatistic (solid line) for the first three rows of Table 3.

[9] In practice follow Hausman and Newey (1995) and take the weighting function

$$\omega\left(z_1\right) = \begin{cases} 1 & \text{if } z_i \leq q^{95}\left(\left(z_1 - \mu_z\right)' \Sigma_{Z_1}\left(z_1 - \mu_z\right)\right) \\ 0 & \text{if } z_i > q^{95}\left(\left(z_1 - \mu_z\right)' \Sigma_{Z_1}\left(z_1 - \mu_z\right)\right) \end{cases},$$

where $q^{95}\left(\left(z_1 - \mu_z\right)' \Sigma_{Z_1}\left(z_1 - \mu_z\right)\right)$ is the 95%-quantile of $\left(z_1 - \mu_z\right)' \Sigma_{Z_1}\left(z_1 - \mu_z\right)$.

Restricted Model	Excluded information	Test statistic	95% Critical Value	P value
income, age, education	sector, area, occupation	2.0538	1.5192	0
income, age, education	sector, area	2.0508	1.738	0.012
income, age, education	sector	0.5354	1.639	0.928
income, age, education, occupation	sector, area	0.9683	1.6971	0.46
income, age, education, occupation	area	1.0081	1.6439	0.42
income, age, education, occupation	sector	0.2608	1.5775	0.98
income, age, occupation	sector, area, education	1.0429	1.4427	0.404
income, age, occupation	area, education	0.9536	1.726	0.49
income, age, occupation	education	0.2211	1.8184	0.97

Note: The distribution of the teststatistic is obtained by the wild bootstrap as in Haerdle and Mammen (1993). We use 250 bootstrap iterations. We normalize the mean of the teststatistic to unity.

Table 3: Which information is contained in individual information sets?

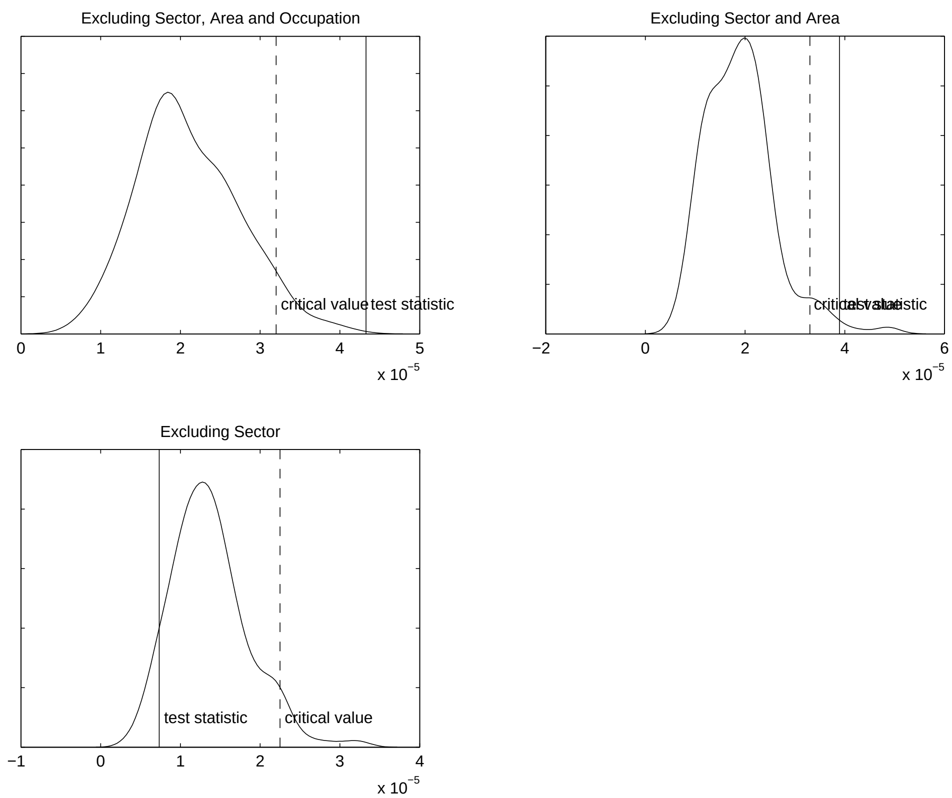

Figure 1: The Content of Information Sets: Distribution of the Teststatistic

4 The Value of Information

The fact consumers seems to exclude information from their information set, which we as econometricians would include, is consistent with the presence of information processing costs as stressed in the literature on rational inattention Sims (2003); Mackowiak and Wiederholt (2000); Luo (2008). In this section, we are going to quantify such processing costs within the realms of a standard life-cycle model of consumption. This model is not only a very natural starting point to analyze the value of information, but it also follows very natural from our econometric application: predicting future income if precisely the the crucial forecasting problem, individuals have to perform.

Our approach is the following. We consider a standard life-cycle problem, where individual face income risk and markets are exogenously incomplete in that only a risk-less bond is available. There are no other constraints on borrowing. To parametrize the income process, requires us to distinguish between the predictable component of future income and the perceived innovation. It is at this point, where differences in the agents' information set come in. Given the same microdata on income realizations, variations in the information set used to predict future income growth, will lead to different decompositions of the income process into its predictable and unpredictable components and to different behavior as encapsulated in the policy function. To estimate the willingness to pay for information, we will therefore first solve for the optimal consumption and savings policies under the individuals' information set. We will then simulate life-cycle profiles using these policy functions but having income evolve under the law of motion, which we as econometricians could infer from the data. These simulated life-cycle profiles allow us to estimate the utility loss of "operating" under a misspecified information set and the willingness to pay for the econometricians' information set.

4.1 The Environment

We consider a parametrization of the life-cycle model that is standard in the literature and for example used by Carroll (1997); Gourinchas and Parker (2002) and Deaton (1991). An infinitely lived consumer chooses consumption to maximize expected utility

$$U = E\left[\sum_{t=1}^{\infty} \beta^t u\left(C_t\right)\right], \tag{31}$$

subject to the per-period budget constraint

$$A_{t+1} = R\left(A_t + Y_t - C_t\right), \tag{32}$$

where Y_t denotes personal income at time t and A_t are the individuals' savings between t and $t+1$. Given an initial condition A_0 and the No-Ponzi condition, (31) and (32) fully characterize the agents' optimal consumption choices for any particular income process $\{Y_t\}_t$ the consumer perceives. We parametrize $\{Y_t\}_t$ in the standard way as

$$Y_t = P_t V_t, \tag{33}$$

where P_t denotes permanent income and V_t is a transitory income shock. The stochastic process for permanent income is given by

$$P_t = G_t P_{t-1} N_t, \tag{34}$$

where G_t denotes the predictable growth in permanent income and N_t is a shock to permanent income. (33) and (34) provide a very parsimonious parametrization of the income process, which nevertheless has been shown to capture salient features of individual income data reasonably well (see e.g. Gourinchas and Parker (2002)). Individuals only need to know the distribution of shocks V_t and N_t and the predictable growth process $\{G_t\}_t$ to know the entire joint distribution of their income process. In particular, suppose that V_t and N_t were log-normally distributed with parameters (μ_V, σ_V^2) and (μ_N, σ_N^2). Then, $(\mu_V, \sigma_V^2, \mu_N, \sigma_N^2, \{G_t\}_t)$ fully characterizes the income process. The concept of permanent income implies that $E[Y_t|P_t] = P_t$, so that (33) requires $\mu_V = -\frac{1}{2}\sigma_V^2$. Similarly, we can always normalize $\mu_N = -\frac{1}{2}\sigma_N^2$ and adjust G_t accordingly.[10]

How would the agents *in this model* predict $(\sigma_V^2, \sigma_N^2, \{G_t\}_t)$? We assume that they follow the rationale of econometricians and hence follow the approach laid out in Carroll and Samwick (1997). Letting $y_t \equiv ln(Y_t)$ (and for the other variables analogously), the growth rate of income is given by

$$y_{t+1} - y_t = p_{t+1} + v_{t+1} - p_t - v_t = g_{t+1} + n_{t+1} + v_{t+1} - v_t. \tag{35}$$

Similarly, the h-step difference is

$$r_{h,t} \equiv y_{t+h} - y_t = \sum_{m=1}^{h} g_{t+m} + \sum_{m=1}^{h} n_{t+m} + v_{t+h}^i - v_t^i. \tag{36}$$

According to the logic of the model, g_{t+1} is the predictable component of income growth, i.e. given *their* information set, the agents would estimate

$$E\left[y_{t+1} - y_t|\mathcal{F}^I\right] = g_{t+1} - \frac{1}{2}\sigma_N^2. \tag{37}$$

From (36) and (37), individuals could then calculate the residual

$$\omega_{h,t} \equiv r_{h,t} - \sum_{m=1}^{h} E\left[y_{t+m} - y_{t+m-1}|\mathcal{F}^I\right] = \sum_{m=1}^{h}\left(n_{t+m} + \frac{1}{2}\sigma_N^2\right) + v_{t+h}^i - v_t^i \sim \mathcal{N}\left(0, h\sigma_N^2 + 2\sigma_V^2\right).$$

[10]Suppose the true process has $ln(N_t) \sim \mathcal{N}(\mu, \sigma^2)$. Then $ln(P_t) | \sim \mathcal{N}(g_t + \mu + p_{t-1}, \sigma_N^2)$. As μ is known to the agent, we can always incorporate in the predictable component g_t and normalize $\mu_N = -\frac{1}{2}\sigma_N^2$.

$\mathcal{F}$	σ_V^2	σ_N^2	$E[g_t]$
age, education	0.056	0.0139	0.0343
age, occupation	0.0557	0.0143	0.0346
age, education, occupation	0.0554	0.0143	0.0345
age, education, occupation, sector	0.0552	0.0141	0.0345
age, education, occupation, sector, area	0.0547	0.0143	0.0345

Table 4: Perceived income processes as a function of the information set

Hence, given more than 2 observations of income (i.e. a sufficiently long panel), σ_N^2 and σ_V^2 can be estimated from $\left\{\omega_{h,t}^2\right\}_h$.

It is clearly seen from (37) how differences in the information set $\mathcal{F}^I$ will lead to different interpretations of the same data $\{y_{i,t}\}_{i,t}$. Not only will the predictable component of income growth be different, but the backed out residual $\omega_{h,t}$ will also have different statistical properties, which will lead the decision maker to arrive at different estimates for the variance of transitory and permanent shocks.

Table 4.1 below reports the results of this exercise for different information sets. Consider for example the first row. If individuals were to use only age and educational status to forecast their income growth rate, they would conclude that transitory shocks had a variance of 0.056 and permanent shocks one of 0.0139. If they were also include their occupational status (row three), their perceived transitory uncertainty would slightly decline, while their perceived permanent uncertainty would slightly increase. Hence: Which information individuals use to distinguish between predictable income growth and noise is important for the conclusions they draw. However, Table 4.1 also shows that the the differences induced by variations in the information set are not very large. In how far these differences in agents' model environment translate into utility difference is the subject of the next section.

4.2 Optimal Consumption Behavior

Our empirical results indicate that individuals' information sets are well described by including only age, occupational characteristics and education. Hence, the relevant data to solve the life-cycle problem is contained in row three of Table 4.1, which together with (31), (32), (33) and (34) fully describe the individuals' problem. As usual, it is convenient to write the problem recursively. Conditional on permanent income P_t, the only additional state variable is cash-on-

Parameter	Value
β	0.96
R	1.02
θ	2
$(G, \sigma_V^2, \sigma_N^2)$	see Table 4.1

Table 5: Parameter values for life-cycle problem

hand $X_t = A_t + Y_t$. This yields the recursive formulation

$$\begin{aligned} V(X,P) &= \max_{A'} \left\{ u\left(X - \frac{1}{R}A'\right) + \beta E^I \left[V(X',P')\,|P\right] \right\} \\ \text{s.t.}\quad & X' = A' + Y' \\ & Y' = GPN'V', \end{aligned} \tag{38}$$

where E^I denotes the expectations taken over the perceived joint distribution of N' and V'. As usual we assume that u takes the CRRA form $u(c) = \frac{c^{1-\theta}}{1-\theta}$. Then it is easy to establish that the problem in (38) is homogenous of degree in P. This allows us to express everything relative to permanent income, i.e. for example $x_t = \frac{X_t}{P_t}$. Transforming the problem in that way yields the easier problem

$$v(x) = \max_{a'} \left\{ u\left(x - \frac{1}{R}\alpha'\right) + \beta G^{1-\theta} E^I_{N'V'} \left[N'^{1-\theta} v\left(\left[\frac{a'}{GN'} + V'\right]\right)\right]\right\}, \tag{39}$$

which crucially only has a single state variable. (39) can be solved numerically in a straightforward manner to yield policy functions π_c^I and π_a^I, where the superscript "I" stresses that these policies are contingent on the individuals' information set. To solve this model, we take standard parameter values, which are displayed in Table 4.2 below. The properties of the solution of this problem are well known and summarized in Figure 2 below. As in Carroll (1997), the consumer displays buffer stock behavior. For low value of cash-on-hand (relative to permanent income), the marginal propensity to consume is high (Panel 3) and cash-on-hand will grow on average (Panel 2). Once a "target level" of cash-on-hand is reached, where cash-on-hand is expected to stay constant, the marginal propensity to consume declines substantially and is similar to the one of certainty equivalence consumers for high values of cash-on-hand. In particular, the consumption function is concave, as is the value function.

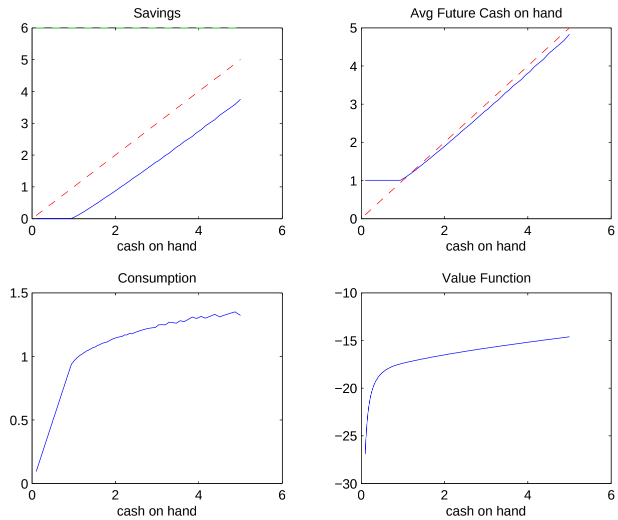

Figure 2: Solution of the lifecycle problem

4.3 The Willingness to Pay for Information

By how much would consumers do better if they were to use a more complete information set? In our empirical analysis above we have seen that individuals' information sets were well described by only containing income, age and occupational characteristics (see Table 3). Comparing rows two and five of Table 4.1 shows the consequences of these coarse information sets - because individuals use too little information, they erroneously assign variations in their income process to transitory shocks, even though such changes could be predicted based on sectoral, educational and regional information. Given this reasoning, we are going to measure to willingness to pay for this improved forecast by the following criterion: how much would a consumer be willing to pay, if she could use the full information set to estimate $(G, \sigma_N^2, \sigma_V^2)$ instead of the one observed in the data.

To answer this question, we are going to adopt the following procedure. Let π_c^I and π_a^I be the policy functions of a consumer with too small an information set and call $\left\{Y_t^I\right\}_t$ the income process implied by this information set. In contrast, let $\left\{Y_t^F\right\}_t$ be the income process under the full information set, i.e. when using all the valuable information to estimate the predictable component of income growth g_t. In our application, this refers to the last row of Table 4.1. Now suppose a consumer were to base his behavior on $\left(\pi_c^I, \pi_a^I\right)$ when facing the income process $\left\{Y_t^F\right\}_t$. How much would he be willing to pay to be able to use the policy functions $\left(\pi_c^F, \pi_a^F\right)$,

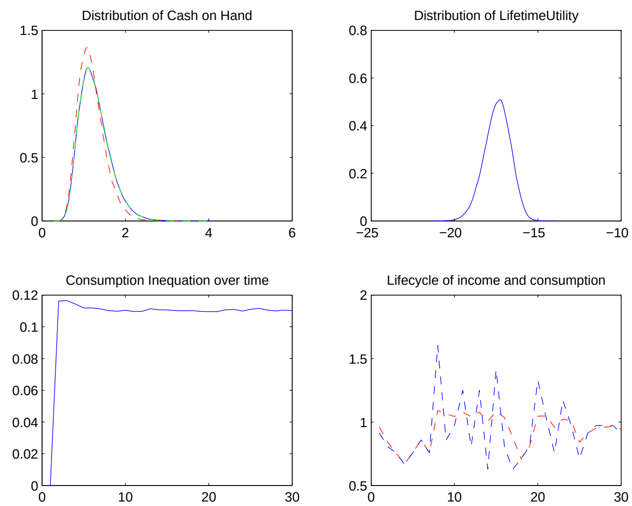

Figure 3: Simulating the lifecycle

which are the solution to the life-cycle problem, when the income process is indeed perceived to be $\left\{Y_t^F\right\}_t$. We think of this willingness to pay is a lower bound on agents' information processing costs.

To calculate these welfare losses numerically, we are simulating M life-cycle profiles using the income process $\left\{Y_t^F\right\}_t$, but behavior based on $\left(\pi_c^I, \pi_a^I\right)$.[11] Hence: consumers face an income process, which has slightly *less* transitory uncertainty than they thought when they made their consumption and savings plans. With N and V being both entirely idiosyncratic shocks, this corresponds exactly to the empirical distribution of future histories, a consumer could experience. Figure 3 below contains some characteristics of this exercise. Panel 1 plots the stationary distribution of cash-on-hand as the solid blue line. As expected from the policy functions in Figure 2 it is centered around the "target-level" of cash-on-hand, where cash-on-hand is expected to remain constant. Panel 2 plots the distribution of utilities, one particular of the M histories induce. The ex-ante value is of course simply the average of this distribution. Panel 3 shows the evolution of the cross-sectional consumption inequality, which is increasing as long as the distribution of cash-on-hand "fans out" and is stable otherwise. Finally, column 4 shows a typical time-series of consumption and income. As expected: consumption is considerably smoother than income as holdings of cash-on-hand provide a buffer for income shocks.

[11]In practice we take $M = 50.000$.

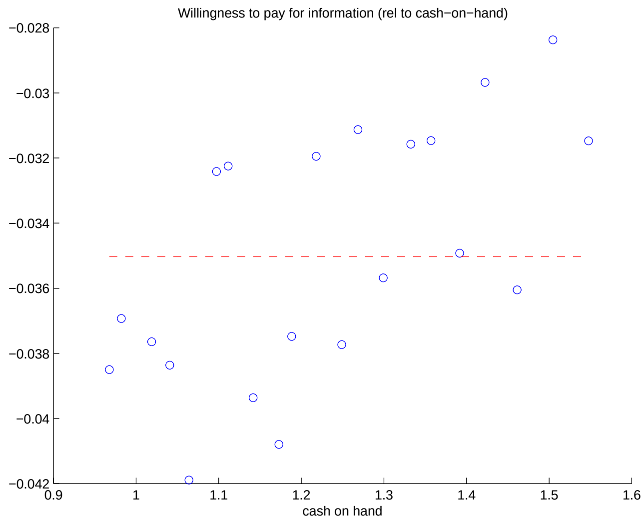

Figure 4: The willingness to pay for information

To measure the willingness to pay for superior information, we can simply redo this analysis for behavior based $\left(\pi_c^F, \pi_a^F\right)$, i.e. for the policy functions derived under the correct income process. The difference in ex-ante values of these two scenarios is exactly the utility loss of using a coarse information set. Formally, let $V_F^I(x)$ and $V_F^F(x)$ be the value of facing the income process $\left\{Y_t^F\right\}_t$ with behavior governed by $\left(\pi_c^I, \pi_a^I\right)$ and $\left(\pi_c^F, \pi_a^F\right)$ at a level of cash-on-hand x. We then define the willingness to pay for information $\Delta^{I,F}(x)$ implicitly by

$$V_F^F\left(x\left(1+\Delta^{I,F}(x)\right)\right) = V_F^I(x). \tag{40}$$

Hence, $\Delta^{I,F}(x)$ is the required relative change in cash-on-hand, which would make an informed consumer equally well off as the less informed consumer. By construction we should have $V_F^I(x) < V_F^F(x)$ so that $\Delta^{I,F}(x) < 0$. We also expect $\Delta^{I,F}(x)$ to be increasing in x, i.e. rich consumers should be hurt less by the misperception of their income process - after all, they are cash-rich, which makes income a less important determinant of their well-being. The results of this exercise in our application are contained in Figure 4, which depicts $\Delta^{I,F}(x)$ for the support of the stationary distribution of cash-on-hand. Two things stand out. First of all, the magnitude of the utility loss form coarse information is relatively small. On average, consumers would be willing to pay 0.035% of their cash-on-hand (relative to permanent income). Given that the median cash-on-hand in the stationary distribution is around 1.2 and that average earnings, which can proxy for permanent income, in Italy are around $23000, this implies that the average willingness to pay for superior information is around \$8.50 (= \$23000*1.2*0.0003). Hence, the utility loss is very minor, which could be expected from the fact that the only source of imperfect behavior is the slight overestimate of the variance of transitory shocks. Secondly, while there is some noise, the willingness to pay for information is higher for poor households as expected by the intuition of buffer-stock savings behavior.

5 Conclusion

What information do individuals use when they form expectations about future events? In this paper we present an econometric framework to answer that question and apply our methods to the case of individuals' income expectations. Using micro-data on agents' beliefs about wage growth, we show that information sets are relatively coarse: while individuals do incorporate occupational characteristics and their age (or their labor market experience) in their income forecasts, we do not find evidence for educational characteristics, sectoral affiliation or local labor market conditions to matter. As this information is self-reported, i.e. in principle available, we interpret this informational coarseness as being consistent with costly information processing. To gauge the utility consequences of this behavior, we calibrate a standard consumption life-cycle model using consumers' information sets from the micro-data. On average consumers would be willing to pay 0.05% of their permanent income to have access to the information set of the econometrician. This represents a lower bound on the costs of information processing.

References

Attanasio, O and G. Weber, "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence for the Consumer Expenditre Survey," *Journal of Political Economy*, 1995, *103*, 1121–1157.

Blundell, Richard, Luigi Pistaferri, and Ian Preston, "Consumption Inequality and Partial Insurance," *American Economic Review*, 2008, *98* (5), 1887–1921.

Brown, B. W. and S. Matial, "What Do Economists Know? An Empirical Study of Experts' Expectations," *Econometrica*, 1981, *49*, 491–504.

Browning, Martin and Annamaria Lusardi, "Household Saving: Micro Theories and Micro Facts," *Journal of Economic Literature*, 1996, *34* (4), 1797–1855.

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

— , "A Theory of the Consumption Function, With and Without Liquidity Constraints," *Journal of Economic Perspectives*, 2001, *15* (3), 23–45.

Carroll, Christopher D., "Macroeconomic Expectations of Households and Professions Forecasters," *Quarterly Journal of Economics*, 2003, *118* (1), 269–298.

— **and Andrew A. Samwick**, "The Nature of Precautionary Wealth," *Journal of Monetary Economics*, 1997, *40* (41-71).

Cunha, Flavio, James Heckman, and Salvador Navarro, "Separating uncertainty from heterogeneity in life cylce earnings," *Oxford Economic Papers*, 2005, *57*, 191–261.

Deaton, Angus, "Saving and Liquidity Constraints," *Econometrica*, 1991, *59*, 1221–48.

Dominitz, Jeff, "Earnings Expectations, Revisions, and Realizations," *The Review of Economics and Statistics*, 1998, *80* (3), 374–389.

— **and Charles F. Manski**, "Using Expectations Data to Study Sobjective Income Expectations," *Journal of the American Statistical Association*, 1997, *92*, 855–867.

Gourinchas, Pierre-Olivier and Jonathan Parker, "Consumption over the Life-Cycle," *Econometrica*, 2002, *70* (1), 47–89.

Guiso, Luigi, Tulio Jappelli, and Daniele Terlizzese, "Income Risk, Borrowing Constraints, and Portfolio Choice," *American Economic Review*, 1996, *86* (1), 158–172.

Guvenen, Fatih, "Learning Your Earning: Are Labor Income Shocks Really Very Persistent?," *American Economic Review*, 2007, *97* (3), 687–712.

Haerdle, Wolfgang and Enno Mammen, "Comparing Nonparametric Versus Parametric Regression Fits," *The Annals of Statistics*, 1993, *21* (4), 1926–1947.

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

— **and Frederic Mishkin**, "The Sensitivity of Consumption to Transitory Income: Evidence From Panel Data on Households," *Econometrica*, 1982, *50* (2), 461–81.

Hausman, Jerry and Whitney Newey, "Nonparametric Estimation of Exact Consumper Surplus and Deadweight Loss," *Econometrica*, 1995, *63* (6), 1445–1476.

Hurd, M. D. and K. McGarry, "Evaluation of Subjective Probabilities of Mortality in the Health and Retirement Study," *Journal of Human Ressources*, 1995, *30*, S268–S292.

Jappelli, Tullio and Luigi Pistaferri, "Using subjective income expectations to test for excess sensitivity of consumption to predicted income growth," *European Economic Review*, 2000, *44* (2), 337–358.

Keane, Michael P. and David E. Runkle, "Testing the Rationality of Price Forecasts: New Evidence from Panel Data," *American Economic Review*, 1990, *80*, 714–735.

Krueger, Dirk and Fabrizio Perri, "How Do Households Respond to Income Shocks?," 2011. Working Paper.

Lovell, Michael C., "Tests of the Rational Expectations Hypothesis," *American Economic Review*, 1986, *76*, 110–124.

Luo, Yulei, "Consumption Dynamic Under Information Processing Constraints," *Review of Economic Dynamics*, 2008, *11*, 366–385.

Machina, Mark and Clive W. J. Granger, "Forecasting and Decision Theory," in Graham Elliott, Clive Granger, and Allan Timmermann, eds., *Handbook of Economic Forecasting*, Vol. 1, Elsevier Science Publishers, 2006.

Mackowiak, Bartosz and Mirko Wiederholt, "Optimal Sticky Prices under Rational Inattention," *American Economic Review*, 2000, *99* (3), 769–803.

Manski, Charles F., "Measuring Expectations," *Econometrica*, 2004, *72* (5), 1329–1376.

Pesaran, M. H. and Martin Weale, "Survey Expectations," in Graham Elliott, Clive W. J. Granger, and Allan Timmermann, eds., *Handbook of Economic Forecasting*, Vol. 1, Elsevier Science Publishers, 2006, chapter 14, pp. 715–776.

Sims, Christopher A., "Implications of Rational Inattention," *Journal of Monetary Economics*, 2003, *50* (3), 665–690.