We are grateful to a very large number of people for a number of different reasons. Our thinking about the issues discussed in this paper has been particularly influenced by a set of people, several of whom have been co-authors in several projects. They include: Rob Alessie, James Banks, Richard Blundell, Martin Browning, Angus Deaton, Hamish Low, Tom MaCurdy, Costas Meghir and Luigi Pistaferri. We have discussed many of the issues covered in this paper (and sometimes disagreed) with them. We certainly learned a lot from them. We are very grateful to three referees for comments and suggestions, and to the editor, for comments, suggestions, and incredible patience! Finally, the second author thanks ESRI, Cabinet Office, Tokyo, for hospitality, and many useful comments. The first author's research was partly financed by the ESRC Professorial Fellowship RES-051-27-0135.
Structure of the paper.

1. Introduction and motivation

2. Facts
   2.1 Average Individual behaviour
   2.2 The evolution of the cross sectional evolution of consumption

3. The life-cycle model
   3.1 Preferences
   3.2 Estimating preference parameters
   3.3 Liquidity constraints as an explanation of excess sensitivity
   3.4 Explaining income tracking and the retirement consumption drop
   3.5 Evidence from the levels of consumption
   3.6 Simulation results

4. Budget constraints and Markets: theory and evidence
   4.1 Complete Markets
   4.2 Exogenously incomplete markets
   4.3 Endogenously incomplete markets
   4.3.1 Imperfect information
   4.3.2 Imperfect enforceability of contracts

5. Alternative models
   5.1 Some puzzles
   5.2 Modifying the basic model
   5.2.1 Relaxing geometric discounting
   5.2.2 Relaxing intertemporal separability
   5.2.3 Financial literacy and information

6. Conclusions
1. Introduction.

In the early 1950s, the prevailing model of consumption behaviour used by macro-economists was inspired by the ‘fundamental psychological law’ mentioned by Keynes (1936) in the General Theory. At that time, the theoretical and empirical limitations of that model became increasingly clear. From a theoretical perspective, it is difficult to construct coherent models based on intertemporal optimizing behaviour that are consistent with Keynes’ description of the ‘fundamental psychological law’. From an empirical point of view, it seemed that Keynes’ view was inconsistent with a number of facts, both at the macro and the micro level. At the aggregate level, for instance, it was observed that the marginal propensity to consume out of disposable income was lower in the short run than in the long run. In cross sections, on the other hand, saving rates seemed to change systematically with the level of income. Moreover, it was observed that groups of individuals with, on average, lower level of income (such as blacks) had higher saving rates than other groups with higher levels of average income (such as whites) at any income level. Finally, it was observed that saving rates are systematically related to changes in income, being higher for individuals experiencing income increases and lower for individuals experiencing income decreases (see Katona, 1949).

All these observations clearly contradicted the implications of the Keynesian model and led to the formulation of the life cycle and permanent income models (Modigliani and Brumberg, 1954, 1980 and Friedman, 1957). These models combined theoretical consistency, in that intertemporal consumption and saving choices were set within a coherent optimization problem, with the ability of fitting most of the facts mentioned in the previous paragraph. The saving rates of blacks was (and is) higher than that of whites, at any income level, because the permanent income of blacks is lower and therefore, conditioning on a common income level, one selects the blacks with higher level of temporary shocks that should, according to the model, be saved. Similarly, individuals with income increases are more likely to be affected by positive transitory shocks. At the macro level, short run fluctuations in disposable income are more likely to be dominated by the variance of temporary shocks that would be averaged out in the long run. Some of these facts still hold in modern data, as we document in Section 2.

The development of the ideas in the seminal contributions of Modigliani and Blumberg and Friedman also led to the realization of other implications. In a simple version of the life cycle model, if income is hump shaped and declines at retirement, consumers will save when they are
young to support consumption in the last part of life and dis-save when they are old.

Modigliani and Blumberg then showed that this fact can explain the correlation between aggregate growth and aggregate saving: growth implies that, in a given year, younger cohorts, who are saving, are ‘richer’ in life-time terms, than older ones, who are dis-saving. The higher the rate of growth is, the larger the difference in resources between savers and dis-savers and, therefore, the higher the aggregate rate of saving.

After its initial development, the other important step in the development of the life-cycle/permanent income model, which is currently used as the standard workhorse of modern macroeconomics, was a rigorous treatment of uncertainty. In the late 1970s, the contributions of Hall (1978) (and MaCurdy (1981, 1999) in the context of labour supply) exploited the idea of using the first order conditions of the intertemporal optimization problem faced by the consumer to derive testable implications of the model. This approach, known as the Euler equation approach, makes possible the empirical analysis of a problem that is analytically intractable by circumventing the need to derive closed form solutions. This is achieved by focusing on the economic essence of the model: consumers, at the optimum, will act to keep the marginal utility of wealth constant over time. The marginal utility of wealth is at the same time a sufficient statistic for consumer choices and, given its dynamic properties, can be ‘differenced out’ in a way which is analogous to the treatment of fixed effects in econometrics.

The Euler equation approach became the standard approach as it allowed to both test the validity of the model and to estimate some of the structural parameters of the utility function. A hypothesis that received much attention, since Hall (1978), is that lagged values of income, or predictable changes in income, do not predict future consumption once current consumption is accounted for. Perhaps as a consequence of this focus on testing, when it came to policy analysis and debates the model and in particular the empirical evidence that has been accumulated on it have been rarely used. One of the reasons for this divorce between the literature on the life cycle model and what should have been its practical use in the design and evaluation of public policy stems from the fact that the Euler equation does not deliver a consumption function. While it can be used to test the model and estimate some of its parameters, it cannot be used to determine the effects of specific policy changes on consumption or saving.

At the same time, much of the evidence that came to be perceived as the accepted view, pointed to rejections of the life cycle model that took the form of ‘excess sensitivity’ of consumption to income. Indeed, in the next section, we take this evidence as one of the
starting points of our discussion of the life cycle model, of its empirical plausibility and of its utility for policy analysis. We have two main goals: to take a stand on where the literature is and what the main issues are and to discuss the public policy implications of the life cycle/permanent income models.

The life cycle model can be loosely defined as a framework where individuals maximize utility over time given a set of intertemporal trading opportunities. Even at this level of generality, the model is of some usefulness. It establishes a conceptual framework that treats the intertemporal allocation of resources in a way which is similar to the allocation of resources among different commodities. Decisions will then depend on the total amount of resources (in the intertemporal context: current and future income as well as current wealth), on preferences over the different commodities (in the intertemporal context: present and future consumption, and possibly bequests) and on relative prices (interest rates and intertemporal trade opportunities).

Without being more specific, however, it is not possible to say much more than what stated in the previous paragraph. Or, saying it differently, this level of generality encompasses many different types of behaviour and has almost no testable implications. In what follows, therefore, we construct a specific model and analyze its components. This exercise forces us to make a number of strong assumptions and modelling choices that we discuss below. We choose to work with a version of the model that is flexible enough to be brought in a serious way to the data and that allows us to derive specific implications on a number of policy-relevant questions.

We start our approach by discussing, in Section 2, a number of empirical findings. We refer to both time series and cross sectional findings and we focus especially on results that might point to empirical rejections of the model. We organize our discussion of the empirical evidence in two parts. We first discuss evidence that refers to individual consumption behaviour. We then move on to look at evidence derived from movements in the distribution of consumption, which allows researchers to look at the functioning of markets and the smoothing of various types of shocks.

After reviewing this empirical evidence, in Section 3 we discuss how a relatively standard but sufficiently rich version of the life cycle model can be made consistent with it. Moreover, we discuss the evidence on the size of the relevant structural parameters. Having established that the model is not wildly at variance with the data and some of the evidence that was presented as a rejection of the life cycle model can be reconciled with it if one specifies a version that is
flexible enough, we go ahead and use the model to quantify, by using simulations, its main properties. In particular, we show how consumption changes with changes in income and interest rates for different values of the structural parameters. The use of simulations is necessary in this context because it is not possible to obtain closed form solutions. Simulations are also useful to study aspects of life cycle behaviour that cannot be studied with the Euler equation approach (such as durables, housing etc.), because transaction costs lead to infrequent adjustments.

Besides preferences and income processes, the other important component of the life cycle model is the intertemporal budget constraint. A specific hypothesis about the nature of the intertemporal budget constraint implicitly assumes a certain market structure and the instruments consumer have to move resources over time (and across states of the world). Section 4, therefore, is devoted to the discussion of alternative market structures, starting from the benchmark of complete markets to move on to various models of incomplete markets.

One of the themes of the paper and in particular of Section 3 is that one can construct rich versions of the life cycle of the model that are not inconsistent with some aspects of the micro data and can be useful in the conduct of policy analysis. Having said that, it is clear that the simplest versions of the model are inconsistent with various aspects of the data and that the empirical literature on consumption has accumulated a number of puzzles. In Section 5, we discuss some of these puzzles and possible extensions and modifications of the basic model. Section 6 concludes the paper.

2. Facts.

In this section, we present some well known facts about consumption behaviour both at the aggregate and at the micro level. Our aim is to present empirical evidence that is or might be relevant to judge the validity of the life cycle model. Indeed, many of the facts that we list below were presented as explicit tests of the life cycle/permanent income model and sometimes interpreted as rejections of the model. In addition to these facts, however, we will also report some new evidence on old findings that motivated the development of the life cycle model.

We divide the empirical evidence we present in two parts. We first discuss findings that refer to individual behaviour. In this first subsection, we consider how individual consumption moves,
on average\(^1\). We then move on to facts about the cross sectional dispersion of consumption and interpret movements in time of these moments as informative about risk sharing and insurance markets available to individuals.

2.1. Average Individual behaviour.

As was mentioned in the introduction, the life-cycle/ permanent income model was developed to explain some facts about consumption. Some of these facts were noticed in aggregate statistics: (non-durable) consumption expenditure is less volatile than income and the marginal propensity to consume seems to be smaller in the short run than in the long run. These ‘macro facts’ still hold and some can also be found in micro data (such as the relative variability of non durable consumption and income - see Attanasio, 2000, and Attanasio and Borella, 2006). Other facts explicitly mentioned by the seminal contributions that originated the life cycle/ permanent income model emerged from cross sectional studies and in particular from observations of how saving rates vary in the cross section with income. As with the ‘macro’ facts, these empirical regularities still hold in recent data. If one looks at US Consumer Expenditure Survey (CEX) data, one finds that the saving rate of blacks is higher than that of whites at any income level, as noted by Friedman (1957). Similar evidence can be obtained in US and UK if one looks at the saving rates by current income level of other groups that differ by the level of ‘permanent’ income, such as households headed by individuals with different levels of education. Analogously, if one considers separately individuals whose income has increased and individuals whose income has decreased, the saving rate of the latter is smaller than that of the former, as noted fifty years ago by Modigliani and Brumberg (1954), citing work by Margaret G. Reid.

The fact that these empirical regularities still hold is important and we come back to them when discussing the empirical validity of the life cycle model. At this stage, we simply stress that the life cycle/ permanent income model offers a coherent explanation for them. The main ideas behind the use of the life cycle model to explain these facts, is that consumers have concave utility functions and therefore prefer smooth paths of consumption (over time and across states of the world) over variable ones. Therefore, only unanticipated changes in income that are perceived as permanent will induce substantive changes in consumption. Expected and temporary changes to income should not induce a strong change in consumption. The

---

\(^1\) Which moment is considered to represent the measure of location of the distribution of individual consumption is an interesting issue which we discuss in what follows.
explanation of the facts mentioned above boils down to the observation that a large fraction of
the changes in income considered in these stylized facts are temporary. For instance, if one
classifies individuals with different levels of permanent income by the level of current income,
one will find that, for each current income level, individuals from the group with a lower level
of permanent income will have a higher level of temporary income, which, the model suggests,
should be saved.

Interestingly, the empirical criticisms of the life cycle model that have been accumulating since
have mainly pointed out deviations from the prediction that expected changes in income
should not be incorporated into consumption. These deviations can be classified into three
groups: those that identify correlations between expected changes in income and consumption
at low frequencies, those that consider short-run fluctuations linked to changes in earnings and
income and those that refer to short-run fluctuations that are linked to ad-hoc payments not
necessarily related to labour supply behaviour.

Low frequency, life-cycle patterns. Carroll and Summers (1991), in an influential paper, show that life
cycle profiles of income and consumption track each other. For many countries both income
and consumption life cycle profiles are hump shaped, in that they increase during the first part
of the life cycle to reach a peak a few years before retirement and decline afterwards. Groups
and countries that exhibit relatively ‘steep’ income profiles also exhibit relatively ‘steep’
consumption profiles. Carroll and Summers, therefore, conclude that income and consumption
track each other over the life cycle, therefore contradicting one of the main predictions of the
life cycle model.

We reproduce this type of graph in Figure 1, where we report life cycle profiles for disposable
income and non durable consumption for two education groups in the UK (the FES data used
here cover the 1978-2007 sample period). We thus adopt the same methodology as Carroll and
Summers (1991). The message that comes out of these pictures is very similar to theirs: at life
cycle frequencies, consumption profiles do follow income profiles. (This is even more strikingly
true if total expenditure replaces non-durable consumption).

A drawback with this type of graphs is that they average over individuals by age, irrespective of
their year of birth. If different generations have access to different life-cycle resources (as
assumed in the life-cycle model) this is not the right thing to do. In figure 2, we show what
happens when the data are grouped in year of birth cohorts – and averages are then taken by
age. (In the figure, cohorts are 10-year wide). There is still evidence of income tracking, even
though this is now less clear cut.
Do these pictures constitute a fundamental rejection of the life cycle model? In the next section, we will be arguing formally that the answer is no, both in theory and in practice. Here we simply point out that if one wants to be serious about bringing the life cycle model to the data, one cannot take the simplest version, which is used for pedagogical reasons, but has to take into account that, in all likelihood, consumption needs evolve over time as family composition changes. This argument is made by Attanasio and Browning (1995).

The simplest way to start considering this type of issues is to look at life cycle profiles for consumption that take into account changes in needs, by considering consumption per capita or consumption per adult equivalent, rather than total household consumption. Figure 3 reproduces Figure 2, but using consumption per adult equivalent. As can be noticed the profiles for consumption are now much flatter. We come back to these pictures and to the interpretation of this evidence in what follows.

---

Figure 1 Average income and (non-durable) consumption by education

---

2 We are grateful to Cormac O’Dea for his help with the FES data.
Figure 2 Average income and consumption by cohort and education
Arguably the largest predictable change in income is the one that occurs at retirement: earnings decline considerably as individuals exit the labour force and such decline should be anticipated. An obvious prediction of the life-cycle model of Modigliani and Brumberg (1954) is that individuals, who should have accumulated wealth (either in private assets or in entitlements to pension benefits) should start de-cumulating it to keep a level of consumption consistent to the one afforded before retirement. Hamermesh (1984) was the first to argue that consumers apparently do not save enough to achieve this aim. If households enter retirement with inadequate savings, they must cut their consumption level, contrary to the life-cycle model predictions.

The recent literature has focused on estimating how consumption levels change around retirement. The existence of a consumption fall around retirement is documented for the UK (Banks, Blundell and Tanner, 1998), for the US (Bernheim, Skinner and Weinberg, 2001), and for Italy (Battistin, Brugiavini, Rettore and Weber, 2009) and has come to be known as the retirement consumption puzzle (or retirement savings puzzle). Banks, Blundell and Tanner (1998) find that for ages between 60 and 67, the level of consumption is lower than that...
predicted by a version of the life cycle model by as much as 1.5% on an annual basis. The cumulated consumption shortfall over this age band, where most people retire, is around 10%. For the US, Bernheim, Skinner and Weinberg (2001) estimate a median drop of 14%, but higher drops for low wealth, low income replacement households. They conclude that "31% of the sample reduce their consumption by at least 35 percentage points". Battistin, Brugiavini, Rettore and Weber (2009), who use Italian data, estimate at 9.8 percent the part of the non-durable consumption drop that is associated with retirement (food expenditure falls instead by 14%).

**Business Cycle Frequency.** The evidence mentioned so far refers to a relationship between predictable changes in income and consumption at the life cycle frequency. Many papers have also looked at the relationship at higher frequencies. This work is typically based on the Euler equations that we will be discussing in the next section, but basically tests the hypothesis that, conditional on current consumption, future consumption is not affected by predicted changes in income, or current level of income. This prediction is obviously related to the observations made by the early proponents of the life cycle/permanent income hypothesis between the lack of strong correlation between changes in consumption and income both in cross sections and in the time series. Many studies in the 1980s, instead, found strong rejections of this prediction. Campbell and Mankiw (1989), in one of the best known and cited papers, found that regressing changes in aggregate US log consumption on interest rates and changes in log disposable income, the latter variable attracted a coefficient of 0.4, statistically different from zero, even after instrumenting current variables with lagged ones to avoid picking up the effects of innovations to the level of permanent income. Cambpell and Mankiw (1991) replicate the evidence for the US for a variety of other countries and attribute such a result to the presence of a large number of consumers who follow a ‘rule of thumb’ and set their consumption equal or proportional to their income.

Hall and Mishkin (1982), perform a similar exercise, but using micro data from the US. Using data on food consumption from the PSID, they find a significant correlation between changes in food consumption and lagged changes in income. They interpret this evidence as indicating that about 20% of households set consumption on the basis of current income, rather than following the life cycle model. Another study that has uses micro data is by Zeldes (1989). He uses the same data as Hall and Mishkin (1982), but distinguishes between consumers with low level of assets and high level of assets and finds that the consumption for the former group is more linked to income than the consumption of the latter. Zeldes (1989) explicitly refers to the
possibility that some consumers are affected by liquidity constraints and restrictions to borrowing that do not allow them to set current consumption at the desired level. We come back to the issue of liquidity constraints in the next section.

The evidence mentioned so far is relevant for the life cycle model as it exploits the implications of the theoretical framework for changes in consumption. In the next section we map directly this evidence on the theoretical framework of the life cycle model. However, it is also possible, albeit more complicated, to derive implications of some version of the model for the level of consumption. Intuitively, the theoretical framework implies that innovations to permanent income should be fully incorporated in consumption, while innovations to transitory consumption of income should not. Therefore, if one specifies a time series model of consumption and income and identifies the permanent innovations to the latter variable, the model predicts that these innovations should be translated one to one into consumption. This implies cross equation parametric restrictions on the VAR representation that can be estimated. Campbell and Deaton (1989) pointed out these restrictions and, using aggregate time series data, found that consumption seems to be too smooth in that it does not react sufficiently to innovations to the permanent component of income. Similar findings were obtained by West (1988), Gali (1991) and Hansen Roberds and Sargent (1991). Perhaps surprisingly, no similar test on micro data was performed until the recent paper by Attanasio and Pavoni (2009), who also find ‘excess smoothness’.

Predicted changes in income. The changes in income that we have considered so far are large predictable changes that occur over the life cycle and/or changes that are likely to be related to changes in labour supply. In recent years, a small literature has developed that studies how consumption varies in relation to changes in income that are not only predictable, but also driven by events that do not have any implications for hours worked or labour force participation. In particular, a large number of papers have looked at the effects of tax refunds or other changes linked to administrative issues. Papers in this literature include Souleles (1999), Parker (1999), Hsieh (2003), Browning and Collado (2001) and Stephens (2007). Souleles, Parker, Stephens and, in part, Hsieh find that consumption reacts to changes in the level of resources available to consumers that are fully predictable. Browning and Collado, on the other hand, as well as the second part of Hsieh’s paper, find that consumers do not

3 We are abstracting here from the possibility of insuring permanent shocks and implicitly considering a consumer who has access to a fairly limited portfolio of assets to move resources over time and across states of the world.
4 An exception is Deaton (1992b).
respond to such predictable changes in resources. We come back to the interpretation of these results later.

2.2. The evolution of the cross sectional evolution of consumption

In the previous subsection, we have listed a number of ‘facts’ that have been discussed in the literature on the empirical implications of the life cycle model. All of the evidence there referred to the properties of consumption levels and consumption changes, on average (either by looking at aggregate data or, in the case of individual data to regressions aimed at identifying the behaviour of the average consumer). The evolution of the cross section distribution of consumption – and income –, however, can also be very informative about the relevant model that describes the data.

One of the first papers to notice the implications of a simple version of the life cycle model for the evolution of consumption inequality was Deaton and Paxson (1994). These authors notice that if income has a unit root, in a basic life cycle model, the cross sectional section of consumption increases over time. One can then consider how the cross sectional variance of consumption for a cohort of individuals born in the same year should increase over time as these individuals age. Testing this forecast for the UK, the US and Taiwan, Deaton and Paxson (1994) show that this is effectively the case. As innovations accumulate, the cross sectional distribution of consumption fans out with age. 5

Battistin, Blundell and Lewbell (2009) use a similar argument to explain a remarkable empirical regularity: the cross sectional distribution of consumption seems to be extremely well approximated by a log normal. This is true across a wide variety of countries. Under a standard version of the life cycle model, at any age, (log) consumption is given by past (log) consumption plus a term that reflects an innovation to permanent income. Therefore, by recursive substitution, one gets that log consumption is given by the sum of innovations from the beginning of life to the current age. By the central limit theorem, the sum of independent

---

5 Using repeated cross sectional data or longitudinal data, one can follow the evolution of consumption inequality for any given cohort and estimate how it evolves with age and time. The identification of an average ‘age profile’ for the variance of consumption that is common for different time periods and different cohorts, is complicated by the fact that age, time and cohort are obviously linked and, without additional restrictions or structure, it is not possible to identify separately age, cohort and time effects. Deaton and Paxson (1994) assume some restrictions on time effects. A forthcoming issue of the Review of Economic Dynamics contains a collection of papers from different countries (including the US and the UK) that undertake similar exercises. The shape of the age profile in the US seems to depend crucially on whether one considers total household consumption or consumption per adult equivalent and which adult equivalence schemes are used.
innovation converges to a normal distribution under some regularity assumptions, even if the individual innovations are not normally distributed.

The facts about the evolution of the cross sectional inequality of consumption and income are also used in another study by Blundell and Preston (1998). Under a specific market assumption, they show that the relative evolution of consumption and income inequality can be used to identify permanent and transitory income variances. The idea is relatively simple: if consumers face a simple asset market structure, changes in the variance of the permanent component of income will induce an equal increase in the cross sectional variance of consumption. Therefore the difference between the increase in the cross sectional variance of income and that of consumption will identify the changes in the cross sectional variance of transitory income.

The caveat about the market structure in the last paragraph makes it clear that there is a stringent relationship between the type of insurance markets agents have access to and the evolution of consumption inequality. Given an initial distribution of consumption (however determined) in the presence of perfect risk sharing, that distribution should stay constant (with some technical caveats we will discuss in Section 4). Deaton and Paxson (1994) noticed that in a footnote and presented evidence on the evolution of the cross sectional variance of consumption as a rejection of the complete market model. In an ingenious paper, Jappelli and Pistaferri (2006) exploit that idea by looking explicitly at movements in the relative ranking in the consumption distribution in an Italian survey. As with other papers, they reject strongly the assumption of perfect risk sharing.

Similarly, Attanasio and Davis (1996) by looking at the evolution of relative consumption across different education groups, and relating that to changes in relative wage changes interpret the evidence of a strong correlation at low frequencies between these two variables as evidence against the complete market hypothesis. Interestingly, Attanasio and Davis (1996) cannot reject the hypothesis that at relatively high frequencies (like one year) there is no relationship between consumption and relative wage changes. This seems to indicate that, somehow, at high frequencies wage shocks are absorbed and not reflected in consumption.

Until the early 1990s, as reported also by Blundell, Pistaferri and Preston (2008), consumption inequality has increased substantially, mirroring the increases in inequality in wages and earnings. After the early 1990s, however, the picture is less clear. Krueger and Perri (2009) report that the overall cross sectional variance of consumption in the US has not increased much. Attanasio, Battistin and Ichimura (2007), instead, find that the cross sectional inequality of consumption does increase even in the more recent period. Even though both papers use
the US Consumer Expenditure Survey (CEX), it turns out that the main difference in the results of these two papers stems from the data used. The CEX is made of two independent samples: one, called the interview survey, in which households are asked retrospective questions about their consumption in the quarter preceding the interview, while the other, the diary survey, in which households are asked to keep a diary for two weeks. It turns out that fact that Krueger and Perri use data from the interview survey, while Attanasio, Battistin and Ichimura integrate data from the two surveys, following the practice of the Bureau of Labour Statistics, which uses the diary survey for some commodities and the interview survey for others.

The different evidence about the evolution of consumption inequality in the US emerging from two different components of the same survey, which is also the main sources of information on consumption at the micro level in the largest industrialized country in the world, justifies a small digression about the quality of consumption data. Information about expenditure and even more so about consumption is notoriously difficult to collect in developed countries. At the same time, the importance of this information cannot be understated. Reliable information on consumption is key for a host of issues, ranging from the construction of price indexes, which are used to index a variety of payments, to the assessment of living conditions and the measurement of poverty, to the estimation of different models of individual behaviour and, ultimately, to the design of public policy. And yet, the resources spent in the collection of reliable consumption data are remarkably small. The CEX is a relatively small survey whose quality is perceived to have been deteriorating over the years. While there are signs that data collection in developed countries has become harder as people seem less willing to respond to survey questions, a redesign and improvement of consumption surveys is, in our opinion, very important.

3. The Life-Cycle Model.

In the first part of the previous section, we mentioned a number of ‘facts’, relating to both individual and aggregate consumption. After a brief mention of the facts that motivated the development of the life cycle model (and that still hold in recent datasets), we discussed several

---

6 If one aggregates the CEX using the appropriate weights, one obtains only a fraction of aggregate Personal Consumption expenditure as measured in the National Accounts. Moreover, this fraction has been declining considerably.
facts that could be cast as criticisms of the model, in that they contradict some simple implications of the theory. To summarize, some of these facts are:

1) The age profile of consumption is hump shaped, apparently tracking the age profile of income for each education group; moreover, groups of individuals that have ‘steep’ income age profiles, seem to have steep consumption age profiles;

2) Consumption drops at retirement;

3) The growth rate of consumption seems ‘too’ sensitive to predictable changes in income;

4) Consumption seems to react to changes in available resources that are fully predictable and transitory, such as tax refunds.

In this section, we present the life cycle model in its modern form, and discuss to what extent it provides an explanation for the facts listed above. Facts that go under the first three headings, will be explained by the consideration that the model does not predict that individuals smooth their consumption but their marginal utility from consumption. We leave to the end of this section our interpretation of the facts under the fourth heading.

The main idea of the life cycle model is a very general one: it can be stated by saying that consumers are supposed to allocate resources over time in order to maximize lifetime utility subject to a resource constraint. At this level of generality, the model does not have much empirical content and is not particularly useful. To bring it to bear on data and make it potentially falsifiable, we need to put a bit more structure on its various components. In particular, we have to specify the individual preferences that inform the maximization problem, the nature of the processes generating the resources available to consumers and the type of markets they have access to. In this section, we specify a basic life cycle model with an eye to the features that would help us to explain some of the facts we mention above. In addition, we also discuss how a version of the model that does fit the available data can be characterized and used in a variety of contexts. In Section 4, we discuss the implications for the model and its applications of the facts about the distribution of consumption discussed in the second part of Section 2.

### 3.1. Preferences

The version of the model we consider is one in which a consumer unit maximizes expected utility over a finite interval subject to a set of constraints
\[
Max E_t \sum_{j=0}^{T-t_j} \beta_{t+j} U(C_{t+j}, z_{t+j}, \nu_{t+j})
\]

s.t.

\[
W_{t+j} = W_{t+j} (1 + R_{t+j}^*) + y_{t+j} - C_{t+j}
\]

\[
W_{t+j} = \sum_{i=1}^{N} A_{i,t+j}^j
\]

\[
R_{t+j}^* = \sum_{i=1}^{N} \omega_i^j \rho_{i,t+j}^j + \rho_{i,t}^j R_{i,t+j}^j
\]

\[
W_{t+j} \geq 0
\]

where \(C\) stands for “consumption”, \(z\) for a potentially large vector of observable variables that affect utility (that may be chosen by the consumer, or given to her – this will normally include household composition variables), and \(v\) for unobservable factors also affecting utility. As we shall see demographics play a key role in explaining the way consumption varies with age, particularly in pre-retirement years. We let the discount factor \(\beta\) be time varying to take into account mortality risk (that helps explain why consumption falls in old age - the survival probability falls with age, and this makes the consumer progressively more impatient).

Throughout the paper we neglect the issue of how decisions are taken within the household, and simply assume the household behaves as a unit.\(^7\)

The first constraint is a generic budget constraint where net worth appears together with its return, income and consumption. Some or all components of income can be simultaneously determined with consumption. For instance, it is possible that income is given by the wage rate times the number of hours worked, where the number of hours is one of the components of \(z\). Equations (3) and (4) define net worth, \(W\), and its return - \(\omega_{t+j}^*\) are the portfolio shares (or weights). The return on net worth is given by the weighted average of the individual returns, \(R_{i,t+j}^*\). We assume these returns do not depend on the net position taken by the consumer on each of these assets, \(A_{i,t+j}^j\).

\(^7\) In the collective model of decision making, households are normally assumed to select efficient allocations, as suggested in Chiappori (1988) – see Vermeulen (2002) for a survey of this in a static setting. Browning (2000) is the first paper to look at the implications of relaxing the unitary model assumptions on intertemporal decisions. Mazzocco (2007) tackles the more general problem of household decision making in a T-period uncertain world, by deriving the Euler equations for individual and household consumption. He looks at the case where individuals can commit to future allocations of resources, and where commitment is instead not possible – because separation and divorce are a possible way out.
Equation (5) gives the limit for total net worth at period T. The consumer has to die without debt, that is, she has to pay back her debt with probability one. This simple restriction imposes quantitatively important limitations to the ability to smooth consumption. Suppose, for instance that the income process is not bounded away from zero and can actually take the value zero with some positive (small) probability. If we further assume that the marginal utility of consumption tends to infinity at very low levels of consumption, then the consumer will never want to borrow in such a situation. This is because the presence of debt together with the non-bankruptcy constraint and the possibility that income takes the value of zero would imply assigning positive probability to zero or even negative consumption, which the consumer deeply dislikes. The consumer will then never want to borrow, even small amounts. One can generalize this to situations where the income process is bounded away from zero. In this case, the consumer will not want to borrow more than the present value of the lowest level of income. Similar considerations apply whenever the survival probability is less than one, if longevity risks cannot be fully insured.

A number of important restrictions are assumed in this formulation. First, the consumer is assumed to maximize expected utility. This is a strong assumption which is often used in the literature. Sometimes the Von Neumann- Morgenstern framework is replaced with different axiomatic structures, such as the Kreps-Porteus axiomatization as parametrized by Epstein and Zin (1989, 1991). Second, we are assuming that preferences are additively separable over time. This precludes the consideration of various types of non-separability, ranging from durables to habit formation. We return to this issue below. Third, we are implicitly assuming that it is possible to write down utility as a function of a single commodity. This practice presupposes an aggregation theorem of the type studied by Gorman (1959).

The problem formulated above is able to encompass different versions of the model that have been considered in the literature. In particular, we treat as special cases the standard permanent income/ life cycle model with quadratic preferences, the so-called buffer stock saving as well as flexible versions of the model (with an important role for demographics and labour supply) that have been fitted to the data.

---

8 Expected utility forces a negative relation between risk aversion and intertemporal substitution, but these are two distinct concepts. This prompted Epstein and Zin (1989) to propose an alternative model that is based on Kreps and Porteus (1978) preferences. Unlike expected utility optimizers, Kreps and Porteus consumers care about the time when uncertainty is resolved, even if they cannot take any action as a result. Epstein and Zin (1989) derive a full set of first order conditions – and show that the Euler equation involves not only consumption growth and the interest rate, but also the return on the market portfolio. Epstein and Zin (1991) and Attanasio and Weber (1989) present estimates of the Euler equation for this type of preferences.
We shall show that the flexible versions of the model can indeed explain the first three stylised facts presented at the beginning of the section. In particular, we shall show that the hump in the age profile of consumption is due to the interplay of demographics and prudence, the excess sensitivity of consumption growth to income growth is due to the dependence of the marginal utility of consumption on leisure, while the retirement consumption drop is due partly to adverse shocks inducing retirement, partly to more efficient shopping that is made convenient by the increased leisure time.

In order to prove all this, we need to work out the solution to the optimization problem. Some features of the solution can be understood by looking at the first order conditions, others require the derivation of the consumption function, either analytically (in some special cases) or numerically.

Let us start with a case where the consumption function can be derived analytically. Let utility be quadratic in consumption (and additively separable in its other arguments $z$), and assume that at least one financial asset is freely traded and yields a fixed real return, equal to the constant time preference parameter $\frac{1-\beta}{\beta}$. The first order condition with respect to consumption, or Euler equation, implies that consumption is a random walk:

$$E(C_{t+1} | I_t) = C_t$$  \hspace{1cm} (6)

where $I_t$ denotes information available at time $t$ (Hall, 1978). If consumers have rational expectations, then:

$$C_{t+1} = C_t + \varepsilon_{t+1} \hspace{1cm} E(\varepsilon_{t+1} | W_t) = 0$$  \hspace{1cm} (7)

for all variables $W$ known at time $t$. Equation (7) can be used to derive a consumption function, in the case where no other asset is available to the consumer (as in Bewley, 1977) and the only stochastic variable is labour income. Substituting (7) into the budget constraints, Flavin (1991) shows that consumption is set equal to permanent income, defined as the interest rate times the present value of current and expected future incomes:

$$C_t = \frac{r}{1+r} A_t + \frac{r}{1+r} \sum_{k=0}^{\infty} E(y_{t+k} | I_t)$$  \hspace{1cm} (8)

Equation (8) is derived for the special case of infinite life, but an extension to finite life can be derived.
In this model, the first difference in consumption, or the error term in (7), equals the present value of income revisions, due to the accrual of new information between periods $t$ and $(t+1)$:

$$
\Delta C_{t+1} = \frac{r}{1+r} \sum_{k=0}^{\infty} \frac{1}{(1+r)^k} [E(y_{t+k+1}) - E(y_{t+k+1} | I_t)]
$$

Equation (7) highlights the consumption smoothing properties of the solution, emphasized in the seminal paper by Modigliani and Brumberg (1954). Equation (8) makes clear the other main implication of the model, that was first stressed in Friedman (1957): consumption depends on the present discounted value of future expected income. The interest rate plays the important role of converting future resources to present ones and therefore constitutes an important determinant of consumption. Equation (8) imposes cross equation restrictions on the joint time series process for income and consumption, as noted in Sargent (1978). Equation (9) implies that in appraising the effects of a given policy, for instance a tax reform that affects disposable income, a distinction must be drawn between permanent and temporary changes (Blinder and Deaton, 1985, Poterba, 1988). Another implication of (9) is that saving predicts future changes in income, the so-called ‘saving for a rainy day’ motive (Campbell, 1987).

Quadratic utility implies certainty equivalence: the consumption function (8) is the same as under certainty, once expectations are replaced by realizations. This is convenient for analytical purposes, but clearly restrictive, for instance in its treatment of financial decisions: quadratic preferences imply increasing absolute risk aversion in consumption (or wealth), something that is unappealing on theoretical grounds and strongly counterfactual (riskier portfolios are normally held by wealthier households). Quadratic preferences also imply that the willingness to substitute over time is a decreasing function of consumption: poor consumers should react much more to interest rate changes than rich consumers, after allowance has been made for the wealth/income effect.

The alternative adopted in much of the literature has been to assume power utility and to allow for the existence of a number of risky financial assets. Power utility, also known as iso-elastic, or CRRA utility, is defined as $U(c) = \frac{c^{1-\gamma} - 1}{1-\gamma}$; it converges to $\ln(C)$ for $\gamma = 1$. 

20
Once one deviates from quadratic utility, however, and/or allows for stochastic interest rates, one loses the ability to obtain a closed form solution for consumption. Many of the studies that made this choice, therefore, have focussed on the Euler equations derived from the maximization problem faced by the consumer. The basic first order conditions used in this literature are:

\[ U_{ct} = \lambda_t \]  

(10)

\[ \lambda_t = E[\lambda_{t+1}\beta(1+r^{k}_{t+1})|I_t] \]  

(11)

where equation (11) is valid as long as the k-th asset can be freely traded by consumers.

Equation (10) says that, at each point in time, the marginal utility of consumption equals the Lagrange multiplier associated with the budget constraint relevant for that period, which is sometimes referred to as the marginal utility of wealth. The second condition, equation (11), that is derived from intertemporal optimality, dictates the evolution of the marginal utility of wealth. An equation of this type has to hold for each asset k for which the consumer is not at a corner. This is because the consumer is exploiting that particular intertemporal margin.

The attractiveness of Euler equations is that one can be agnostic about the stochastic environment faced by the consumer, the time horizon, the possible presence of a bequest motive, the presence of imperfections in financial markets (as long as there is at least one asset that the consumer can freely trade), and the presence of frictions in other variables affecting utility, z. All relevant information is summarized in the level of the marginal utility of wealth. The approach is conceptually similar to the use of an (unobservable) fixed effect in econometrics. By taking first differences, one eliminates the unobservable marginal utility of wealth and is left only with the innovations to equation (11). This approach has played an important role in the empirical analysis of the life-cycle model and we will come back to it.

The derivation of a closed form solution for consumption when certainty equivalence does not hold is possible in the case where the utility function exhibits constant absolute risk aversion. Caballero (1991) shows that in a modified Flavin model (with certain finite life and CARA preferences) the optimal consumption age profile is flat with no uncertainty, but increasing with income uncertainty. This change in the slope of the consumption profile is labelled as
precautionary saving, because early in life consumers save more if labour income is more uncertain. Later work by Gollier (1995) and Carroll and Kimball (1996) established that a similar result holds whenever the third derivative of the utility function is positive, and this feature of preferences is labelled prudence. Both CARA and power utility exhibit prudence. The presence and size of precautionary savings is a matter of great relevance for public policy, in so far as public insurance schemes covering such risks as unemployment, health and longevity should reduce the need for consumers to accumulate assets.

The great merit of even this simple model with prudence is that it highlights the need to save for rainy days even if sunny days are equally important. An increased variance in the shocks to income reduces consumption even if expected income does not change. In the case of discrete variables, such as unemployment or illness, changes in first and second moments occur simultaneously, but this is not the case for continuous variables. The ability to distinguish between first and second moments effects is of crucial importance in the analysis of public policy, because public policy can be used to provide social insurance, by reducing the variance while keeping the mean constant. For instance, a revenue-neutral tax reform that cuts taxes for the rich may depress consumption because it induces more precautionary saving (Varian, 1980, stresses the insurance role of a progressive income tax).

3.2. Estimating preference parameters

The Euler equation is particularly useful from an empirical point of view because it can be cast as a set of orthogonality conditions that should hold in a variety of situations and allows estimating preference parameters and testing the validity of the model without being explicit about all the details of the stochastic environment faced by the consumer and without having to solve explicitly the dynamic optimization problem for consumption or other variables jointly determined with consumption. As stressed by Chamberlain (1984), estimation of the Euler equation requires observations covering a long period of time, as the orthogonality conditions hold in expectation, and (but for the special case of complete markets) sample expectations converge to population expectations only over time (see also Hayashi, 1987).

A version of the Euler equation holds even if the consumer chooses labour supply, durable consumption and many other variables that are subject to different types of adjustment costs and frictions. It holds under a wide variety of assumptions about the information set used by
the consumer and, by the law of iterated expectations, it holds whenever the information set used by the econometrician is no larger than that available to the consumer. To use it, one does not need to specify assumptions about pension systems, future wage processes, bequests motives and so on and so forth. Moreover, it reflects the main essence of the life cycle model: the fact that consumption is chosen so to keep (discounted, expected) marginal utility constant over time.

The Euler equation can be used for two purposes: testing for the validity of some of the model assumptions, notably the ability of consumers to save in response to changes in intertemporal prices, and estimating preference parameters. The first paper to estimate a consumption Euler equation (Hall, 1978) was entirely devoted to testing the model, but much of the literature since has done both.

Hall took the case of quadratic utility and a fixed interest rate such that \( (1+r)\beta = 1 \). Under these conditions, equation (6) obtains and preference parameters are not identified. Another notable feature of Hall’s version of the Euler equation for consumption is that it aggregates perfectly, because it involves linear transformations of the data, and can therefore be empirically implemented in micro and aggregate data alike. The Euler equation (6) implies that no variable known to the consumer at time \( t \) should help predict the change in consumption between \( t \) and \( (t+1) \) – an important and easy to test implication of the intertemporal optimization model that has been rejected a number of times on aggregate and micro data alike (Jappelli and Pagano, 1989, Hall and Mishkin, 1982).

The special features of the Hall’s model may explain these rejections – for this reason in the literature Euler equations have been estimated and tested for more general preference specifications. As mentioned earlier, a popular preference specification is the power utility function, given by \( U(c) = \frac{C^{1-\gamma} - 1}{1 - \gamma} \), that has been used in the consumption literature since the papers by Hansen and Singleton (1982 and 1983). Its main advantage is analytic convenience, as it yields first order conditions that are log-linear in consumption. However, such a specification also imposes strong restrictions on preferences. The elasticity of intertemporal substitution of consumption is, in this context, constant and equal to \( 1/\gamma \). This implies that the degree of intertemporal substitutability of consumption is independent of the level of consumption, even at very low levels of consumption. Moreover, the same parameter governs both the elasticity of intertemporal substitution and the degree of risk aversion. This is the
consequence of the assumption of intertemporal separability and separability across states of
the world.
Substituting equation (10) into (11) and using the properties of the power utility function the
Euler equations for consumption corresponding to each asset (k) are:

\[
E_{it} \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} \beta(1 + r_{t+1}^k) = 1
\]

(12)

Where \( \gamma \) is a curvature parameter (equal to the relative risk aversion parameter and to the
reciprocal of the elasticity of intertemporal substitution) and \( \beta \), the subjective discount factor,
measures patience. Equation (12) is an orthogonality condition stating that a particular
transformation of the data is orthogonal to the information set used by the agent. Such a
condition suggests naturally the use of some GMM method to estimate the unknown
parameters and, to the extent one considers a vector of variables whose dimension is greater
than that of the parameter to be estimated, to test the validity of the model. In essence, Hall
(1978) was the first test, in a specific context, of this orthogonality condition.
An equation such as (12) can be log-linearized to obtain (see Hansen and Singleton 1983):

\[
\Delta \ln C_{t+1} = \alpha_{t+1} + \frac{1}{\gamma} \ln(1 + r_{t+1}^k) + \epsilon_{t+1}^k
\]

(13)

where \( \alpha_{t+1} \) is a time-varying term that depends on the preference parameters \( \gamma \) and \( \beta \) as well on
the conditional second moment of the argument of the expected utility operator in equation
(12).
Estimating equation (12) seems preferable because no assumption has to be made about the
conditional variance term, but will produce inconsistent estimates whenever there is
measurement error in consumption. Equation (13), instead, can be consistently estimated if
there is serially uncorrelated measurement error, as long as one can find instruments that are
orthogonal to both the error term and the time varying intercept. Attanasio and Low (2004)
discuss conditions under which equation (13), estimated under the assumption of a constant
\( \alpha_{t+1} \), yields consistent estimates for the curvature parameter, \( \gamma \). Notice that the other preference
parameter, the discount factor, is not identified in this framework, as it gets buried into the constant.\footnote{Equations (12) and its log-linearized version (13) refer to an individual asset. If the consumer has access to several assets for which she is not at a corner, one can consider an Euler equations for each of these assets. These equations have been used extensively to study the implications of the model we are considering for asset pricing, since Lucas (1978) and Breeden (1979).}

Of particular importance for policy analysis is $1/\gamma$, or elasticity of intertemporal substitution, that tells us how the marginal rate of substitution between today and tomorrow’s consumption reacts to changes in the interest rate, keeping life-time utility constant. The increase in the interest rate represents a decrease in the price of future consumption relative to current consumption, and this induces a `substitution effect’ of a decrease in current consumption and a commensurate increase in current saving. This would be counteracted by an `income effect’ since with a higher interest rate a given target level of future consumption is achieved with less saving. As noted by Summers (1981), wealth effects, concerning the amount that expected future incomes are discounted, reinforce substitution effects and also lead to a decrease in consumption or increase in saving when the interest rate goes up. These wealth effects tend to be stronger when the time period that the individual cares about is longer. Ultimately, which of these forces dominates depends on preference parameters and is, therefore, an empirical issue, that depends on the size of the elasticity of intertemporal substitution.

An influential paper by Hall (1988) claimed that this parameter is close to zero. This finding has been challenged on various grounds. A low response of consumption growth to the real interest rate could obtain if some consumers are liquidity constrained, or if the error term correlates with that part of the real interest rate that is explained by the instruments. Attanasio and Weber (1993, 1995) point out that aggregation bias could be responsible for such low estimate: the aggregate consumption growth rate is computed by taking logs of the mean of individual consumption, whereas equation (13) implies that means of the logs should be taken instead. Attanasio and Weber (1993) provide evidence that the difference between these two terms is highly serially correlated, thus invalidating lagged consumption growth as an instrument. When they correct for this they find higher estimates of the elasticity of intertemporal substitution. Attanasio and Weber use cohort data (that is: data from repeated cross sections, that is consistently aggregated over individuals born in the same years): when they focus on cohorts of individuals who are least likely to be liquidity constrained, and control for changes in taste shifters, they estimate a much higher elasticity (around 0.8) using UK
(1993) and also US cohort data (1995). Recently, Scholz, Seshadri and Khitrtrakun. (2006) address the issue of how well the life cycle model predicts wealth holdings, and take as benchmark case $1/\gamma = 0.33$, but they also show that the model fits best when they take $1/\gamma = 0.67$. In a recent, very ingenious paper, Engelhardt and Kumar (2007) use differences in employer’s matching rates in 401(k)s and its effect on participation to identify the elasticity of intertemporal substitution, and obtain a point estimate of 0.74.

In much of the macro literature, the iso-elastic specification has played a predominant role. Little attention has been paid to the possibility that the elasticity of intertemporal substitution may differ across consumers, particularly as a function of their consumption. A simple way to capture the notion that poor consumers may be less able to smooth consumption across periods and states of nature is to assume that the utility function does not depend on total (non-durable) consumption, rather on the difference between consumption and needs. Thus we could have retained the analytical attraction of power utility, but have $(C-C^*)$ as its argument, where $C^*$ is an absolute minimum that the consumer must reach in each and every period. This functional form is known as Stone-Geary utility in demand analysis (see Deaton and Muellbauer, 1980, chapter 3, for example), and is the simplest way to introduce non-homotheticity in a demand system.\footnote{One could interpret “external habits” (Abel, 1990, Campbell and Cochrane, 1999) as a special way to parameterize $C^*$ (by making it a fraction of past consumption).} Attanasio and Browning (1995) take a different route and extend the isoelastic specification by modelling marginal utility as a quadratic function in the logarithm of consumption. Blundell, Browning and Meghir (1991), Atkenson and Ogaki (1996) and Guvenen (2006) are among the few other examples of papers that explicitly allow for wealth-dependent elasticity of intertemporal substitution (see also Crossley and Low, 2005). However, a recent paper by Battistin, Blundell and Lewbel (2009) suggests that non-durable consumption is log-normally distributed, and this is consistent with the standard iso-elastic utility specification.

3.3. Liquidity constraints as an explanation of excess sensitivity

The Euler equations (12) and (13) have been estimated mostly on aggregate data. In several cases, some of the model implications have been rejected: generally speaking, the error term has been found to correlate with information available at time $t$ (rejection of the over-
identifying restrictions), and in particular with that part of income growth that could be explained by such information (excess sensitivity). A good example of this type of results are those in the influential Campbell and Mankiw (1990 and 1991) papers, which report results from a regression like (13) where changes in aggregate consumption were related to changes in (expected – as instrumented) disposable income. The significance of the expected income coefficient is interpreted in that paper as a fundamental violation of the basic model, caused either by ‘rule of thumb’ consumers, consuming a fixed proportion of their disposable income, or by binding liquidity constraints. In fact, a reason why excess sensitivity or violations of the over-identifying restrictions may be detected is because some consumers are not able to borrow and lend at the same interest rate. Binding liquidity constraints may cause excess sensitivity if constrained individuals experience temporary income changes: they will change consumption by more than the intertemporal optimization problem implies. However, excess sensitivity may also have other explanations, as we shall see later.

Liquidity constraints can take several forms – in the next section we shall consider market structures in which such constraints are the optimal response to information asymmetries or enforceability problems. However, much of the literature imposes such constraints exogenously. If, in addition to the non-bankruptcy constraint considered in the previous section, one imposes some exogenous and more stringent limits on the amount people can borrow, it is possible that consumers will be constrained in a given period and the Euler equation (12) will not hold. In this case, assuming a variable rate of interest \( R_t \) one would have:

\[
E_t \left[ \beta (1 + R_{t+1}) u'(c_{t+1}) / u'(c_t) \right] < 1
\]

The consumer would like to increase current consumption and therefore her current marginal utility is higher than in the case in which the borrowing restriction is not binding. The presence of a binding liquidity constraint means that the consumer is at a kink of the intertemporal budget constraint, so that the tangency requirement between the ratio of marginal utilities and intertemporal prices holds as a slack condition.

The presence of a binding liquidity constraint represents an important issue for the empirical application of the Euler equation. Of course, the borrowing restriction will not be binding in every period and, when not binding, the Euler equation will hold. However, even in periods in which the liquidity constraint does not bind and the Euler equation holds, the level of consumption will be affected as the consumer takes into account the possibility that the constraint will bind in future periods. As pointed out, for instance, by Hayashi (1987), the
presence of a borrowing restriction is equivalent to a shortening of the time horizon: a consumer who expects to face a binding liquidity constraint \( n \) periods ahead, will plan to have zero wealth in that period, therefore behaving as if the planning horizon was \( n \) periods. \(^{11}\) Notice, however, that the relationship between consumption at \( n-1 \) and \( n-2 \) is not affected, and the Euler equation between those two periods holds as if the liquidity constraint is not operative. The liquidity constraint has an effect on the level of consumption even when is not binding. In addition to the extreme case of an exogenously given borrowing limit, one can consider alternative borrowing restrictions. For instance, it is possible to consider the case of a difference between borrowing and lending interest rates, or more generally, the case in which the interest rate varies with the position of the consumer in a given asset, typically increasing with higher levels of debts. These cases have been studied, for instance, by Pissarides (1978) and Juster and Shay (1964).

A direct way to detect binding constraints is to ask consumers whether they applied for and were denied credit. Jappelli (1990) reports that 12.5% of the 1982 wave of the SCF respondents answered they were denied credit, and models the probability of credit denial as a function of observable characteristics. The problem with this type of question is that consumers may have been denied credit for good reasons (likely violation of the no-bankruptcy condition), or may have decided not to apply for credit on the assumption that this would be refused to them (a discouraged borrower effect).

Less direct tests for liquidity constraints that meet these criticisms are based on the idea that the Euler equation should be violated for groups of consumers who are likely to be constrained, such as the young and those whose liquid assets are particularly low. This strategy was implemented by Zeldes (1989) using the ratio of liquid assets to income at time \( t \) as an indicator of potential constraints. Zeldes reports evidence for liquidity constraints among households with very low liquid assets: for this group consumption growth would rise by 4% if the constraint were relaxed. However, any sample split based on choice variables may induce endogenous selection, particularly if the error term reflects preference heterogeneity.

For this reason, other papers have followed a different route, that works whenever the amount borrowed depends on other variables, such as earnings (see Alessie, Melenberg and Weber, 1989, and Weber, 1993) or collateral (such as in the case of durables, Brugiavini and Weber, \(^{11}\) Deaton (1991), simulating a stationary economy with impatient consumers and precautionary saving, shows that liquidity constraints are rarely binding)
1994, Alessie, Devereux and Weber, 1997, and Chah, Ramey and Starr, 1995). In this case the presence of a binding borrowing restriction distorts not only the intertemporal margin but also the allocation of resources between different commodities (or leisure and consumption) within a period. Alessie, Devereux and Weber note that identification of liquidity constraints is greatly enhanced if the relationship between the borrowing limit and the choice variable is exogenously changed within the sample period (this is true in their analysis of cars and non-durable expenditure, because the hire-purchase terms were heavily regulated by the UK government over the first part of the sample period, completely unregulated later). They do find evidence of binding liquidity constraints in some of the years prior to financial liberalisation, but only for young consumers. Finally, a test for liquidity constraints that compares the first order conditions across periods to the first order conditions across goods is proposed by Meghir and Weber (1996) – their results suggest liquidity constraints may be binding only for young consumers.12

Despite all these different approaches, the most widely cited piece of evidence for the operation of liquidity constraints is “excess sensitivity”. But excess sensitivity of consumption to income (both at low and high frequency) may be due to incorrect preference specification, as we argue in the next section.

### 3.4. Explaining income tracking and the retirement consumption drop

An influential paper by Carroll and Summers (1991) uses micro data to document ‘excess sensitivity’ of consumption to income. The authors notice not only that the life cycle profiles of income and consumption track each other, but that the shape of the two profiles co-vary across different groups in the population. For instance, households headed by an individual with low education have a relatively flat profile for both income and consumption, while households headed by better educated individuals present more of a hump shape. This evidence has been

---

12 A problem with all Euler equations-based tests as well as with the direct question is that, as Hayashi (1987) explains, the presence of an operative, albeit not binding liquidity constraint is equivalent to a shortening of the planning horizon. This may be the relevant information that is needed for policy purposes. Evidence on this can be obtained by noting that consumers that are liquidity constraints will not be sensitive to changes in the level of the interest rate. As they will be at a kink of an intertemporal budget constraint, the demand for loans will be inelastic to changes in the slope of such an intertemporal budget constraint: the interest rate. This approach has been followed by Juster and Shay (1965) and Attanasio, Goldberg and Kyriazidou (2008). Interest rate elasticities of credit demand have been estimated by Gross and Souleles (2002) using US credit card data, and by Alessie, Hochguertel and Weber (2005) using Italian instalment credit data.
used to argue that consumers are impatient but prudent, to the point of holding liquid assets to buffer shocks – this has come to be known as the buffer-stock model of savings.

The different results mentioned above are reminiscent of the early debate between Thurow (1969) and Heckman (1974). The former pointed to the covariance over the life cycle between income and consumption as a rejection of the life cycle model, while the latter replied that a version of the life cycle model where consumption and leisure were not separable could well explain such a covariance. The micro papers cited above show that consistent with Heckman (1974)’s argument, excess sensitivity can be reconciled with the intertemporal optimization model if more general, and sensible, utility functions are used. In particular, if one assumes that leisure affects utility in a non-additive way, consumption changes respond to predictable labour income changes, whether or not leisure is a freely chosen variable.

Micro data equations typically show the need to take account of the effects of some time-varying characteristics on preferences: demographics and leisure. A way to introduce this dependence is to specify period-t utility as:

\[ u_t = \left( e^{\delta t} C_t \right)^{1-\gamma} \]

(15)

where \( Z \) contains hours of work, and other taste shifters (Attanasio and Weber, 1993 and 1995), some of which might be unobservable. If one takes the model to micro data one has to allow for the effect that demographic variables have on utility. The fact that ‘when you have a wife and a baby a onepenny bun costs three pence’ (Gorman) has to be taken into account if one estimates the model on micro data. Demographics might explain consumption changes as well as the shape of the consumption age profile, as argued by Browning and EIjnaes (2002). The increase in household size early in life, and decrease past age fifty, can explain why consumption age profiles are hump-shaped, in apparent contradiction of the consumption smoothing implications of the life-cycle theory. The interaction between demographics and prudence explains instead why the peak in consumption occurs later in life than the peak in household size and can generate consumption-income tracking for four different education groups when labour income is uncertain, as shown in Attanasio, Banks, Meghir and Weber (1999).

---

13 Leisure has also been introduced in the utility function in some papers that use aggregate data, such as Mankiw, Rotemberg and Summers (1985), and Bean (1986).
More general preference structures, that allow the elasticity of intertemporal substitution to depend also on current consumption, have been considered in the empirical literature, (Attanasio and Browning, 1993, Blundell, Browning and Meghir, 1994, Attanasio and Weber, 1995, Meghir and Weber, 1996). As in the standard case, the parameters of these specifications can be estimated using the Euler equations and other first order conditions of the optimization problem faced by the consumer.

The results obtained in the papers that use the Euler equation to estimate preference parameters and test the model could be summarized by saying that a flexible version of the life cycle model is not rejected by individual level data, especially if one focuses on households headed by prime aged individuals, that is, excluding very young households and households on the verge of or passed the retirement age. Typically, there is no excess sensitivity of consumption growth to income growth once changes in leisure and demographics are taken into account.

These results show that it is possible to find a specification of preferences that is not inconsistent with the available micro data. However, leisure and demographics variables could capture the essence of the predictability of income and make the estimates of the ‘excess sensitivity’ parameter imprecise. Such variables, according to this interpretation, therefore should not be interpreted as ‘taste shifters’. There are two possible answers to this objection. First, a ‘horse race’ between expected income and these other variables seems to indicate that the introduction of the latter does not just inflate the standard error, but also reduces the size of the income coefficient. Second, once one has estimated the life cycle model augmented with these additional variables, one should ask whether the implied preferences are sensible and predict features of the data that were not used to estimate them. We will come back to this issue below.

Finally, aggregation issues have been proven to be important. As pointed out in Attanasio and Weber (1993), the difference between the consistently aggregated equation (based on the means of the logarithm of consumption) and what is available in macro data (the logarithm of the mean) is a highly persistent process that correlates with lagged information. Attanasio and Weber (1993) also show that results obtained with improperly aggregated micro data are consistent with results obtained with aggregate data and indicate rejections of the model that instead disappear with properly aggregated data and rich enough preference structures.
In models where utility depends on the consumption of several goods and leisure, one can explicitly allow for home production of goods and services (Becker, 1965 and 1981; Ghez and Becker, 1975). This has been recently emphasized in connection with changes in spending behaviour around retirement (Aguiar and Hurst, 2005). The availability of time-use data allows testing for the implications of the model in terms of changes of the composition of consumption over the life-cycle (Aguiar and Hurst, 2007 and 2009).

A number of recent papers have estimated the effects on changes in consumption of well-defined predictable tax changes (such as tax rebates, social security withholding tax), often finding these effects to be different from zero (Parker, 1999, Souleles, 1999, Shapiro and Slemrod, 2003, Johnson, Parker and Souleles, 2006). This violation of the model predictions is surprising, because consumption does not appear to react to other anticipated income changes (Browning and Collado, 2003, Hsieh, 2003).

The evidence of these “natural experiment” papers suggests that consumption reacts to predicted changes in disposable income only to the extent that these changes are relatively small, as noted by Browning and Crossley (2007), because small optimization errors might have trivial utility costs (Cochrane, 1989).

A number of recent papers report evidence in favour of a liquidity constraints interpretation. Stephens (2008) shows that consumption reacts to the repayment of vehicle loans, and this is particularly true for young individuals, who are more likely in principle to be liquidity constrained. Agarwal, Liu and Souleles (2007) investigate credit cardholders’ response to the 2001 tax rebates and find that most people first increased repayments, but then the young and those who were initially close to their credit card limit started spending more (and building up debt faster). The eventual rise in spending could then be attributed to the operation of liquidity constraints. Similarly, Hsieh, Shimizutani and Hori (2008) find that Japanese consumers’ response to a spending coupon program tailored to families with children and the elderly was highest among those with low wealth.

Another piece of evidence that apparently contradicts the life cycle model is the retirement consumption puzzle. The simple life-cycle model of Modigliani and Brumberg (1954) predicts that individuals save during their working lives to keep their consumption level constant once they retire. Hamermesh (1984) was the first paper to argue that consumers apparently do not
save enough to achieve this aim. If households enter retirement with inadequate savings, they
must cut their consumption level, contrary to the life-cycle model predictions.

The recent literature has focused on estimating how consumption levels change around
retirement. The existence of a consumption fall around retirement is documented for the UK
(Banks, Blundell and Tanner, 1998), for the US (Bernheim, Skinner and Weinberg, 2001), and
for Italy (Battistin, Brugiavini, Rettore and Weber, 2009) and is known as the retirement
consumption puzzle (or retirement savings puzzle).

Banks, Blundell and Tanner use British cohort data and show that the standard Euler equation,
in which consumption growth is a function of intertemporal prices and changes in
demographics, over-predicts the level of consumption by as much as 1.5% on an annual basis
for ages between 60 and 67. The cumulated consumption shortfall over this age band, where
most people retire, is around 10%. They argue that only a fraction of this drop can be
attributed to the increased leisure time that accompanies retirement. Later work by Smith
(2006) uses information on food for UK households who retired over the sample period, and
stresses the importance of distinguishing between voluntary and involuntary retirement: a
significant drop for food consumption is observed only for those who retire early because of
poor health or job loss. Indeed, David M. Blau (2008) stresses that consumption drops at
retirement can be reconciled with life-time optimization if there is uncertainty over layoffs, job
offers, health and mortality, and retirement is a discrete event that is freely chosen by the
household. However, in Blau’s model the causal effect of retirement on consumption is zero.

Bernheim, Skinner and Weinberg use PSID data to estimate Euler equations for food
consumption. The retirement status is instrumented by taking age-specific predicted
probabilities conditional on demographics. The sample is split in groups: low wealth-to-income
households drop their consumption most. BSW estimate a median drop of 14%, but higher
drops for low wealth ratio, low income replacement households. BSW conclude that "31% of
the sample reduce their consumption by at least 35 percentage points". The evidence they
provide is consistent with the notion that consumers do indeed enter retirement with
inadequate savings. A number of papers have further investigated the issue on US data - Haider
and Stephens (2007), who estimate a smaller consumption drop for those who retire at the
expected time; Fisher, Johnson, Marchand and Smeeding (2005), who use CEX data, deflate
expenditure by the squared root of household size, and estimate a smaller drop (around 2.5%)
for total expenditure than for food consumption (around 5.7%).
Recent papers by Aguiar and Hurst (2005 and 2007) and Hurd and Rohwedder (2006) stress that the drop in expenditure at retirement does not necessarily imply an increase in the marginal utility of consumption. For instance, worker-related expenditure (transport to and from work, canteen meals and business clothing) is no longer needed - whether they account for a large enough part of pre-retirement consumption is an open issue. Also, home production of services (laundry, gardening, house-cleaning, cooking) may become advantageous, and the extra leisure time may allow consumers to shop more efficiently. This last channel has recently been stressed by Aguiar and Hurst (2005 and 2007) in their careful analysis of food consumption around retirement, whilst the increase in home production of services by recent retirees has been documented by Hurd and Rohwedder (2006), who exploit time-use data. The literature has investigated as further reasons for this drop unexpectedly low pensions or liquidity problems as well as time-inconsistent behaviour (Angeletos, Laibson, Repetto, Tobacman and Weinberg, 2001). Another recent paper by Aguila, Attanasio and Meghir (2010) that looks at changes in consumption around retirement (using the longitudinal dimension of the CEX in the US) find that the decline in food expenditure is compensated by increases in non food items, so that the total is roughly constant.

Battistin, Brugiavini, Rettore and Weber (2009) use Italian data and instrument retirement with public pension eligibility. To be more precise, they take a regression discontinuity approach, and make the identifying assumption that spending behaviour would be smooth around the threshold for pension eligibility if individuals did not retire. They estimate at 9.8 percent the part of the non-durable consumption drop that is associated with retirement induced by eligibility (food expenditure falls instead by 14%). They show that this fall is not driven by liquidity problems for the less well off in the population, and can be accounted for by drops in expenses that are work-related or leisure substitutes. However, they also show that retirement induces a significant drop in the number of grown children living with their parents, and this can account for most of the retirement consumption drop.14

As Hurst (2008) recently put it, we should no longer talk about the retirement consumption puzzle, rather about “the retirement of a consumption puzzle”. Once preferences are correctly modelled, home production is taken into account, and attention is focussed on those who retire

---

14 A few recent papers study how different expenditure items vary over the life-cycle and in relation to leisure. Miniaci, Monfardini and Weber (2002 and 2009) focus on changes around retirement age, while Aguiar and Hurst (2009) look at the evolution of work-related expenses over the whole life-cycle. This last paper claims that, once allowance is made for the effect of changes in family size and composition, all the decline in consumption can indeed be attributed to the fall in work-related expenses.
at the expected age, then the drop in food spending and total spending around retirement does not imply a violation of the model prediction that consumers smooth marginal utility over time.

3.5. Evidence from the levels of consumption

Although consumption (growth) appears on the left hand side of equation (13), that is not a consumption function, but an equilibrium condition. It cannot explain/predict consumption levels, even conditional on current consumption: consumption is crucially determined by the residual term $e_{t-1}^k$ and there is nothing in equation (13) that tells us what determines such a term or how it changes with news about income, interest rates or any other relevant variable, including future ones. This inability to predict how consumption moves in response to changes in the economic environment is the price one pays if one stays within the remit of the Euler equation. What is bought in terms of robustness is paid in terms of the non-availability of a consumption function.

As stressed above, the Euler equation imposes some restrictions on the dynamics of consumption but, on its own, does not determine the level of consumption. Neglecting numerical complications, a solution for consumption can be obtained considering jointly the Euler equation and the sequence of budget constraints faced by the consumer as well as her initial wealth and a terminal condition. As noted by Sargent (1978), Flavin (1981) and later by Campbell (1987), the Euler equation and the intertemporal budget constraint imply a number of cross-equation restrictions for the joint time series processes of consumption and income. When one is able to obtain a closed form solution for consumption, as is the case with quadratic utility, these restrictions can be easily expressed in terms of a linear time series model and tested.

To be more specific, given an intertemporal budget constraint that assumes a fixed interest rate and a relatively general process for labour earnings, with quadratic utility, the level of consumption is given by equation (8). Given an assumption on the time series process for income, this equation will imply cross equation restrictions on the bi-variate time series model for consumption and earnings. Some of these restrictions are implied by the Euler equation, while others are not. In particular, the restrictions on the contemporaneous correlation between income and consumption are not: as we stressed above, the Euler equation is silent about how news about income are translated into news about consumption.
To consider a specific example let us assume a simple AR representation for labour income: $A(L)y_t = a + \zeta_t$, where $\Lambda(L)$ denotes a polynomial in the lag operator. In this case, equation (8) implies (Flavin, 1981) that

$$A\left(\frac{1}{1 + r}\right)\Delta C_{t+1} = \frac{r}{1 + r} \zeta_{t+1}$$

(16)

This relation provides a link between the variance of the income shock, $\zeta_{t+1}$, and the variance in the error term of the Euler equation, $\epsilon_{t+1}$. If there is enough persistence in income growth (positive serial correlation in the first differenced process), then (16) implies that $\Delta C_{t+1} = \psi \epsilon_{t+1}$ with $\psi > 1$, and consumption growth should vary more than income growth over time.

Notice that the conditioning set for the expectations about future earnings in equation (8) is left unspecified. An advantage of the approach pursued here is that one can then condition equation (8) (and the corresponding equation for earnings) on a smaller information set, and obtain a similar expression. This implies that the approach is robust to the presence of an informational advantage of the consumer over the econometrician. The reason is that, by looking at consumption, we are implicitly using the information the consumers has at her disposal.

If we follow Campbell (1987) and define saving as:

$$s_t = \frac{rA_t}{1 + r} + y_t - C_t$$

(17)

we can re-write (16) as:

$$s_t = -\sum_{k=1}^{\infty} (1 + r)^{-k} E (\Delta y_{t+k} | I_t)$$

(18)

Equation (18) shows that individuals should “save for rainy days” (future income falls), and holds (by the law of iterated projections) even if we take expectations conditional on a subset of the information used by economic agents, such as past income and saving.

While Flavin (1981) and Campbell (1987) test the cross equation restriction that arise in the quadratic utility case in the VAR representation of income and consumption, Campbell and Deaton (1989) and West (1988) use the same structure to propose a test that links the innovation to permanent income to consumption. These authors present evidence that aggregate consumption is ‘excessively smooth’ in that it does not react enough to news about income. In particular, because the model for earnings seems to be characterized by a unit root and some additional persistence in the first changes, the model would imply that consumption
changes should reflect the permanent income innovation more than one-to-one. Not only is
the income shock permanent, but it also predicts future, smaller shocks of the same sign. This
implies that over the business cycle consumption should be more volatile than income. But in
actual aggregate data consumption is smoother than income.

By taking the intertemporal budget constraint as a given, Campbell and Deaton (1989) make a
connection between excess sensitivity and excess smoothness. Hansen, Roberds and Sargent
(1991) propose a test of the intertemporal budget constraint (given the martingale behaviour of
consumption implied by the Euler equation) that is shown to be similar to the Campbell and
restrictions implied by the intertemporal budget constraint are and what restrictions can be
tested with time series data. They also consider a number of generalizations, such as habits and
other forms of non separabilities. When discussing endogenous liquidity constraints below, we
argue that the test of the intertemporal budget constraint that Hansen, Roberds and Sargent
propose can be interpreted as a test of market structure.

In situations where preferences are not restricted to functional forms that admit a closed form
solution and one considers more realistic environments, one has to rely on numerical methods
to get the consumption function, as shown in the seminal paper by Deaton (1991). We discuss
the literature on numerical solutions and simulations of the life cycle model in the next
subsection.

A less ambitious but potentially profitable approach, that does not require numerical methods
or incredibly rich data sets, is the estimation of reduced form equations, whose specification is
informed by the life-cycle model. These are particularly useful in situations in which one
analyzes large (and possibly exogenous) changes to some of the likely determinants of
consumption or saving. Such studies can address substantive issues and even test some aspects
of the life cycle model. Examples of studies of this kind include the reaction of consumption
(and saving) to changes in pension entitlements (Attanasio and Brugiavini, 2003, Attanasio and
Rohwedder, 2003, Miniaci and Weber, 1999), to swings in the value of important wealth
components (such as housing, Attanasio and Weber 1994) and to changes in specific taxes, as
discussed above.
3.6. Simulation results

A small literature has developed that numerically solves and simulates the intertemporal consumer problem under uncertainty, starting with an influential contribution by Deaton (1991), who studied a model with power utility and infinite life. Deaton considers the existence of liquidity constraints, and shows that impatient consumers hold limited assets to insure against low income draws. Carroll (1992) instead covers the case of finite lives, and shows that if consumers are sufficiently impatient and their labour income is subject to both permanent and temporary shocks, they set consumption close to income, at least until they are in their forties. The model with impatient consumers under labour income uncertainty has been labelled “the buffer stock model”, because saving is kept to the lowest level compatible with the need to buffer negative income shocks. Carroll’s buffer stock model can provide a rationale for the income tracking of consumption that was highlighted by Carroll and Summers (1991).

Later work by Attanasio, Banks, Meghir and Weber (1999) – refined by Gourinchas and Parker (2002) - clarifies the role played by age-related changes in demographics and the hump-shaped age profile of labour income in generating income tracking for relatively young consumers (as mentioned above, micro data show that financial asset accumulation starts in mid-life). Hubbard, Skinner and Zeldes (1994 and 1995) show instead how precautionary motives interact with the insurance properties of social security in the US.

Attanasio, Banks, Meghir and Weber (1999) is the only paper that thoroughly investigates the interaction between demographics and precautionary savings. It does it for four different education groups – in the analysis education matters because income and demographics age profiles are education specific, and because education can affect patience. Note that it imposes that life ends with certainty at age 70 – it does not let the discount factor change with survival probabilities. Survival probabilities can play an important role in determining the shape of the consumption profile at older ages.
To illustrate the interplay of demographics and uncertainty in shaping the consumption age profile, let us consider one of the figures in Attanasio, Banks, Meghir and Weber (1999), that presents the average of a large number of simulated consumption profiles for a specific education group in the US (high school graduates). In Figure 4, NW panel, we plot average consumption age profile simulations corresponding to the solution for the case where there is income uncertainty and demographics evolve with age (“baseline”). Counterfactual simulations shed light on the role of demographics and uncertainty: in the NE panel, there is income uncertainty but demographics do not change with age; in the SW panel, there is no income uncertainty and demographics evolve with age. Finally, the SE panel presents the previous three curves together.

The key conclusion drawn from this and similar pictures for other education groups is that precautionary savings alone would imply a peak in consumption quite late in life, while demographic needs would make consumption peak relatively early. It is the interplay of these two opposing forces that generates a peak in the early to mid-forties, as observed in the data.
Simulation results are also useful to study how consumption reacts to changes in the interest rate, an important topic for both monetary and fiscal policies. It is well known that changes in the interest rate have both substitution and income effects on consumption, going in opposite directions for a net saver. Summers (1981) stresses the role played by the wealth effect that is induced by applying different discount factors to future flows of labour income. Which of these effects prevails depends on the elasticity of intertemporal substitution as well as on other factors affecting preferences. 15

With the isoelastic preferences discussed above, in a two period model, if $\gamma > 1$ the income effect prevails on the substitution effects so that an increase in the interest rate causes current consumption to increase. When $\gamma < 1$, instead, the substitution effect prevails and current consumption decreases. While these effects are quite clear in a simple model, their quantification in a multi-period framework, in which preferences are affected by a variety of factors which are possibly changing over time, is not simple. The effect of changes in the interest rate will depend not only on standard preferences parameters such as the elasticity of intertemporal substitution and the discount rate but also on their interaction with the evolution of needs, on the shape of the income profile, on the institutional arrangements for pensions and so on and so forth.

To gain a better understanding of how these various factors interact and determine the final effect there is little alternative to using numerical methods that solve the model and simulate life cycle trajectories for a large number of hypothetical consumers to obtain average life cycle profiles for different sets of parameters and different interest rates. These types of exercises are certainly not novel. Summers (1981) and Evans (1983), for instance, present simulation results to characterize the interest rate elasticity of savings. Over the last 15 years, however, the literature on the solution of a life cycle model under uncertainty has made great progress, so that we are now able to solve and simulate realistic and much more sophisticated models under uncertainty. These models normally use parameter values from the empirical studies that have estimated preferences from micro data, possibly using Euler equations that are robust to the exact specification of the stochastic environment faced by consumers.

Attanasio and Wakefield (2008) simulate a life cycle model to understand the importance of the elasticity of intertemporal substitution to determine the size of the reaction of savings to changes in the interest rate. They first simulate a model of a single consumer with an isoelastic

---

15 Summers (1981) claims that both for the wealth effect and for general equilibrium effects the interest rate elasticity of saving is bound to be high. Evans (1983) provides several counterexamples to this claim
utility function (they take the elasticity of intertemporal substitution, EIS, $1/\gamma = 1$ as baseline), no bequest motive and a stochastic income process calibrated on UK data. Figure 5, which is taken from their paper, plots the changes in the life cycle asset profile induced by a change in the interest rate from 2% to 2.5% under different elasticities of intertemporal substitution.

First of all notice how, in all cases, the effect of the interest rate on savings depends on age. One of the implications for the aggregate effects of certain policy changes will depend on the demographic composition of the population. Second, notice that the increase induced by interest rate changes is considerably larger in the baseline case when the elasticity of intertemporal substitution is 1 (and the discount rate is 0.025). When the elasticity of substitution is low, the effect up to age 45 is basically zero.
Figure 6 Average change in assets by age when r changes, with & without family

Figure 6 introduces changing needs over the life cycle (for the baseline case where the EIS is unity). These are calibrated using standard adult equivalent schemes and actual UK data. Utility depends not on total consumption but on consumption per adult equivalent. Notice how the effect of a change in interest rate on saving is reduced, even with an elasticity of intertemporal substitution left at the relatively high level of one. Other papers, such as Attanasio and Weber (1993), Attanasio and Browning (1995), stressed the importance of demographics to explain observed patterns of consumption life cycle profiles and to fit the life cycle model to the data. Figure 6 makes the additional point that the evolution of consumption needs also affects the way consumers respond to changes in intertemporal prices.\(^{16}\)

As the age of household members and family composition evolves, the effect that a given level of consumption has on utility obviously changes. In a sense family composition variables play a role in the dynamic allocation of resources which is analogous to the discount factor and, as they vary dramatically over time, can have quantitatively important implications for the level and shape of asset life cycle profiles.\(^{17}\) Moreover, as we mentioned above, their evolution interacts with preferences to determine the response to specific changes in the environment.

\(^{16}\) In both figures 5 and 6 the change in assets is normalised by dividing through by expected income at age 21.

\(^{17}\) Evans (1983) stresses that differences in discount factors can alter both the shape of asset life cycle profiles and the way in which they respond to changes in the interest rate. This is also discussed by Gruber (2006)
4. Budget constraints and markets: theory and evidence

The implications of any model of intertemporal consumption behaviour depend heavily on the nature of market arrangements through which the agents considered in the model interact and on the type of intertemporal trades agents have access to. This is especially true for the life cycle model. As the focus of the life cycle model is the intertemporal allocation of resources (in an uncertain environment), insurance and credit markets are particularly important.

What is often considered the standard version of the life cycle-permanent income model, such as the one we used above to derive Euler equations, makes some stark assumptions about the nature of intertemporal trades available to agents. Typically, state contingent trades are ignored (hence ruling out insurance) and often the consumer considered in the model is only endowed with a single asset that pays a known and fixed interest rate. On the other hand, this consumer might be allowed to borrow or save without limits (except for a non-bankruptcy constraint). Obviously all these assumptions are very extreme and unrealistic. It is therefore important to consider explicitly several alternative market structures. In what follows, we stress how different market structures might have only limited implications for some of the results typically used in the literature (in particular Euler equations), and how they have important implications for the level of consumption and for its relationship to current income.

We do not discuss the portfolio problems implied by the consideration of several assets simultaneously or the implications that can be derived from the theory for asset pricing. Instead we focus on the implications of different market structures for the intertemporal allocation of consumption.\textsuperscript{18} In particular, we consider four different structures. The first, which constitutes a useful, albeit unrealistic, benchmark, is the case in which the consumer is assumed to have access to a full set of state-contingent Arrow-Debreu securities. The second is the standard case, considered above, in which the consumer is endowed with an exogenously given technology to move resources into and from the future at a given rate. We then move on to introduce some specific limits to the intertemporal trades available to consumers, not only in terms of the type of securities available to them, but also, more specifically, in terms of the amount of resources that can be moved from the future. This is what is normally referred to as

\textsuperscript{18} From a theoretical point of view, the consideration of several assets, some of which might have state contingent returns, and others for which several restrictions are relevant, can be easily added to the model. A useful result in this respect is the fact that as long as the consumer is not a corner for a given asset, one can write down an Euler equation for that particular asset, linking consumption at two points in time. This is true regardless of the presence of imperfections and frictions in other markets.
‘liquidity constraints’. Finally, we consider situations where, rather than being imposed exogenously, specific limits to intertemporal trades arise endogenously from the presence of specific imperfections.

4.1. Complete Markets

In this benchmark case, it is assumed that consumers have access to state contingent securities that pay a certain amount of consumption, which depends on the particular state of the world that is realized at a given time. The implicit assumption behind this framework is that very complex contracts can be written and supported by a completely symmetric information structure among consumers and then enforced perfectly.

The stochastic environment in which agents live can be summarized by a vector of state variables \( \theta \). We will denote with \( \theta^t = \{\theta_1, \theta_2, ..., \theta_t\} \) the history of the state variables up to time \( t \).

We also assume that the sets of all possible histories up to time \( t \) are endowed with a certain probability measure. We write as \( \pi(\theta^t) \) the probability of history \( \theta^t \). The consumer receives resources \( y_i \) that depend on the history of the state variables at time \( t \). The consumer problem can be written as:

\[
\max E \left[ \sum_j \beta^t U(C_i(\theta^t)) \right] \quad \text{s.t.} \quad \sum_j p_j(\theta^t)(C_j(\theta^t) - y_j(\theta^t)) \leq 0, \quad \forall \theta^t
\]  

where \( p(\theta^t) \) is the price of consumption at time \( t \) at history \( \theta^t \). To compute the equilibrium that would prevail in such a situation, one has to consider the problem (19) for each consumer in the economy. The specific type of competitive equilibrium that would prevail will depend on the specifics of individual preferences, income processes and, possibly, initial endowments. The computation of this type of equilibrium is in principle very complex. However, the literature provides some important results that allow one to characterize some key features of the intertemporal allocation of consumption in such a situation.

In particular, one can exploit the fact that, in the absence of externalities and other distortions, competitive equilibrium allocations are Pareto efficient and invoke the first Welfare theorem to describe them as the result of the optimization problem faced by a fictitious social planner that maximizes the weighted average of individual utilities. While the theory is silent about the weights used in this problem (that is on the particular competitive equilibrium that is realized),
the study of this problem is useful in characterizing the intertemporal consumption allocations.

If we assume that the social planner has a technology to transfer aggregate resources to and from the future at a rate $r$, we can then re-write equation (19) as:

$$\text{Max}_{\{A_{t+1}, \theta_i\}_{t=0, t+1 \leq T}} \sum_{i} \phi_i \left[ \sum_{i} \beta^i \sum_{\theta'} \pi' \left(\theta'\right) U \left(C^i_{t'}(\theta')\right) \right]$$

subject to $A_{t+1} \leq (1 + r) A_t + \sum_{i} (C^i_{t'}(\theta') - y^i_{t'}(\theta'))$ (20)

where index $i$ refers to individuals, $\phi_i$ is the Pareto weight given to individual $i$, $A_{t+1}$ is the asset available to society to move resources over time. The social planner achieves a given intrapersonal and intertemporal consumption allocation by a set of transfers among the agents.

Notice that in this formulation the presence of individual savings is not necessary. The social planner could achieve any given allocation implied by a certain amount of individual savings without them and via different sets of interpersonal transfers. The problem in (20) assumes the possibility of aggregate borrowing and the existence of a fixed rate at which resources can be moved over time. Moreover, it also assumes that individuals have homogeneous preferences and discount factors. These assumptions can be relaxed. It is also possible to use more general forms of preferences incorporating, for instance, leisure.

From the first order conditions of problem (20), one can derive some interesting relationships. In particular, one set of first order conditions will be:

$$\phi_i \beta^i U_{C^i_{t'}(\theta')} = \mu(\theta') / \pi'(\theta')$$

Equation (21) states that the marginal utility of consumption for individual $i$, at a given state of the world at time $t$, multiplied by her Pareto weight, is equal to the aggregate constraint multiplier relevant in that state of the world (divided by the probability of that particular history).

Notice that the right-hand side of equation (21) does not depend on the index $i$. This implies that consumption for each individual has to be such that the (discounted) marginal utility multiplied by that person's Pareto weight has to be equal to the right-hand side of equation (21). This implication of perfect risk-sharing means that idiosyncratic risk is diversified and only
aggregate fluctuations determine individual consumption. Individuals can have different consumption levels if Pareto weights are different, but over time the cross section of consumption moves to guarantee that (21) holds for each individual.

An important possible implication of the complete markets, full insurance model is that it allows the construction of a ‘representative’ consumer. That is, an important aggregation theorem holds under the hypothesis that markets are complete and idiosyncratic risk is fully diversified. Aggregate consumption moves as if it was determined by a representative consumer who acts according to the life cycle model subject to an intertemporal budget constraint similar to that in equation (20). This result holds, as showed, for instance, by Townsend (1994), even in the presence of heterogeneous preferences (see also Atkeson and Ogaki, 1996). The preferences of the representative consumer will aggregate individual preferences.

Equation (21) has constituted the basis for much of the empirical tests of perfect insurance. Log linearizing (21) and taking first differences one obtains the simple proposition that changes in marginal utility of consumption should be the same in a cross section of consumers who share risk efficiently

\[
\ln(\beta) + \Delta \ln(U_{C_i(\theta')}) = \Delta \ln(\mu(\theta')) / \pi'(\theta')) \equiv \nu_i
\]

(22)

With a specific assumption about the utility function it is easy to derive from equation (22) an expression that can be brought to data. For instance, with CRRA utility, one gets:

\[
y \Delta \ln(C_i') = -\ln(\beta) + \nu_i
\]

(23)

Notice that the left-hand side is individual consumption growth, while on the right hand side we only have aggregate variables, that is quantities that do not vary in the cross section.\(^{19}\) A possible test of perfect risk sharing, therefore is to add to equation (23) a variable related to the change in resources accruing to household \(i\) in period \(t\) and test the significance of such a coefficient.

\[
\Delta \ln(C_i') = k + \nu_i + \varphi \Delta y_i + \epsilon_i'
\]

(24)

where \(y\) could be household income, or wages, the term \(\nu_i\), the multiplier on the aggregate resources constraint, can be captured by time dummies and \(\epsilon_i'\) represents a regression residual. The test of perfect risk sharing, therefore, will be that the coefficient \(\varphi = 0\). It is worth noting that equation (22), and therefore (23), holds without error as perfect insurance assumes the possibility of writing contracts that determine the allocation of resources under any possible

\(^{19}\) In this example we do not consider preference heterogeneity (for example in discount factors). Townsend (1994) discusses cases with heterogeneity in risk attitudes.
state of the world. The presence of a residual in equation (24), therefore, has to be justified by measurement error in either consumption or $\gamma$. In the latter case, however, one has to take into account the fact that estimates of $\phi$ will be affected by attenuation bias.

This implication of perfect risk sharing, first noticed by Townsend (1994) and tested for the US by Cochrane (1991) and Mace (1991) is very powerful and the empirical tests that can be derived from it very appealing. From a theoretical point of view it captures the fundamental idea of risk sharing, that is that risk is pooled efficiently among the participants in a risk sharing agreement. From an empirical point of view, the strategy proposed by Townsend (1994) is appealing because it allows one to test efficient risk sharing without specifying the entire budget constraint relevant for the individual agents. Agents may use a host of different instruments to achieve efficient intertemporal allocations, including a variety of privately held assets, informal interpersonal transfers, implicit contracts and so on. The Townsend (1994) test looks at the actual allocation of resources, regardless of how it was achieved.

Townsend (1994) applies his test to 3 Indian villages and finds some important rejections of full risk sharing. Cochrane (1991) and Mace (1991), instead, use US data and test the implications of equation (23) using different specifications for the utility function. Cochrane (1991) shows that the growth rates of food consumption do not respond to some shocks (such as strikes, or involuntary moves), but are affected by involuntary job loss and long illness. Perhaps surprisingly, Mace (1991) does not reject the hypothesis of full risk sharing. It has been pointed out, however, that her results could be due to measurement error (see Nelson, 1994). Strong rejections of the perfect insurance hypothesis, instead, are reported by both Hayashi, Altonji and Kotlikoff (1996), and Attanasio and Davis (1996). Attanasio and Davis (1996), in particular, show that while short run changes in relative male wages (across education and cohort groups) do not seem to be related to changes in relative consumption, when one considers lower frequencies, one finds significant effects.

An alternative test of perfect insurance can be obtained by looking at the evolution of cross sectional second moments. Considering again the log-linearization of equation (21) one can easily obtain the result that the cross sectional variance of the log of marginal utilities should be

---

20 Altug and Miller (1990) estimate and test a model of household consumption and labour supply choices with complete markets.

21 Hayashi, Altonji and Kotlikoff (1996) use food consumption data from PSID to test for complete risk-sharing across all households and across households that belong to the same family (dynasty); in both cases they reject the null of complete risk-sharing.
constant over time, under perfect risk sharing, a fact first noted by Deaton and Paxson (1994). Deaton and Paxson (1994) stress the fact that, in many countries, the life cycle profile of consumption inequality seems to be increasing with age, a fact that is consistent with simple versions of the life cycle model with a single asset, but not with complete markets. Since that contribution, other papers have tested this implication (see for instance Attanasio and Jappelli, 2001, and Attanasio and Székely, 2004). Further research that tries to relate consumption inequality to income inequality includes papers by Blundell, Pistaferri and Preston (2004), Attanasio, Battitstin and Ichimura (2007), Krueger and Perri (2006) and Guvenen (2007).

The paper by Blundell, Pistaferri and Preston (2007) is particularly interesting because it decomposes innovations to household income into ‘temporary’ and ‘permanent’ components. These authors consider how changes in the variance of permanent and transitory income components are translated into changes in the variance of consumption and estimate the fraction of permanent and transitory income shocks that are effectively insured. They find that a large fraction of temporary shocks are indeed insured, especially in the case of ‘better off’ households, while most (but not all) permanent shocks seem to be uninsured. We will come back to a possible interpretation of these results below.

One can summarize the evidence on the implications of complete markets saying that the large majority of empirical work in this area points to a sound rejection of the hypothesis of perfect insurance. This is true both in developed and developing countries. In developed countries it seems that low frequency, persistent shocks are not completely insured. However, one can often also reject the hypothesis that these shocks are completely uninsured and therefore fully reflected in consumption. This is not inconsistent with the evidence on ‘the excess smoothness of consumption’ we mentioned in Section 3.5. On the other hand, it seems that transitory shocks are, to a large extent, insured. Given this evidence, therefore, the focus of much of the current research is on models without perfect risk sharing and incomplete markets.

4.2. Exogenously incomplete markets.

Much of the literature on consumption assumes that individuals have a certain number of exogenously given assets that they use to move resources over time. What is common to the set of studies we consider here is that the financial market structure and therefore the type of assets individuals have access to is exogenously given, as in the discussion in Section 3. Some
of these papers consider explicitly the asset prices that clear these markets. However, the type of assets considered is exogenously given.

Many studies have considered a situation in which consumers borrow and lend in an asset whose net supply is zero: in such a situation, individuals with (temporary) positive shocks will lend to individuals with (temporary) negative shocks. Models of this type are often referred to as ‘Bewley’ models, from Bewley (1983), who was among the first to study the competitive equilibrium in a model where individuals try to smooth income fluctuations over time. Other versions of the model also consider the presence of aggregate saving, which can play a role in production, as, for instance, in Aiyagari (1994). Finally, while some versions of the model allow for both borrowing and saving, others prohibit borrowing.

The equilibrium conditions that are relevant in the presence of exogenously missing markets are effectively those we discussed in the first part of Section 3 and that have been used for the empirical work on the life cycle model. The Euler equation (12) in Section 3.2 is the relevant equilibrium condition when individuals can borrow and lend at the interest rate $r^k$ in asset $k$. If consumers are prevented from borrowing and are at the corner in terms of their asset position, the relevant condition becomes the inequality in equation (14). It is worth stressing that, even in the presence of borrowing restrictions, the expression in (14) can hold as equality if the liquidity constraints is not binding between two given time periods. 22

Notice the difference between the Euler equation implied by the Bewley model and the first order condition of the social planner problem. First, the first order condition (12) holds in expectations, while equation (21) holds state by state. Second, there is no individual Euler equation in the full insurance case, but something similar to it. To see this, one can appeal to the aggregation theorem referred to in the last paragraph and derive the Euler equation for the representative consumer. In this sense, as is obvious, the model with complete markets is much more restrictive.

We have already discussed the evidence on Euler equations and liquidity constraints in Section 3. We now move on to the discussion of economies where markets are not complete because of the presence of specific imperfections. Unlike the economies we have considered so far, the market structure and the assets available to an individual consumer are not given, but are determined as an equilibrium outcome.

22 Indeed, if one considers the so called ‘natural’ liquidity constraints that is the present discounted value of the lowest income realization with a positive probability of occurring, under some regularity conditions on the utility function, the constraint will never be binding. But even for lower bounds, such as zero, the liquidity constraint might be binding only occasionally.
4.3. Endogenously incomplete markets

As we mentioned above, there is strong evidence that rejects the hypothesis of complete contingent markets that provide full insurance against idiosyncratic risk. An important theoretical and empirical challenge, therefore, is to construct models in which full risk sharing is not achieved in equilibrium because of the presence of specific imperfections. From a theoretical viewpoint, it is certainly preferable to map the nature of imperfections one considers into particular market structures rather than making more or less ad-hoc assumptions about the nature of markets. From a policy point of view, the fact that the market structure is an equilibrium outcome allows one to take into account the possible effects that given policy interventions have on the nature and extent of private insurance markets. This might be fundamental to evaluate the effects of a given policy.

The literature on endogenously incomplete markets has mainly focused on two types of imperfections: imperfect information and imperfect enforceability of contracts. We discuss these two classes of models in turn.

4.3.1. Imperfect information models

In imperfect information models individuals are assumed to have private information either about their income or about the effort they put in producing income (moral hazard). It is therefore necessary, when looking at an insurance market, to guarantee that in equilibrium individuals are induced to reveal their private information. The constrained efficient allocation of resources can be studied in a way similar to the case of perfect insurance, by looking at the problem solved by a social planner. A problem like (20) has to be supplemented with the incentive compatibility constraints that guarantee the revelation of private information. From a methodological point of view, the characterization of these contracts can become extremely complex, as the transfers through which the social planner redistributes income among private agents can be a function of the entire history of individual past income. A big methodological breakthrough in this literature came with the Thomas and Worrall (1990) and Atkeson and Lucas (1992) papers in which these authors re-wrote the problem in terms of ‘promised utilities’. This technical trick allows a relatively simple characterization of the equilibrium.

23 Other important papers in this literature include Green (1987), Phelan and Townsend (1991).
A result that is particularly important in this literature, first derived by Rogerson (1985), can be useful to study the empirical implications of these models (see Ligon, 1998). In contrast to the permanent income model, where the marginal utility of consumption follows a martingale, here it is the reciprocal of the marginal utility of consumption that has this property. In particular, it can be proven that:

$$E_t \left[ \frac{1}{U'(C_{t+1})} \right] = \frac{1}{U'(C_t)} \quad (25)$$

The intuition behind this results follows from the fact that the problem, in the Atkeson and Lucas (1992) approach, is formulated in terms of promised utilities that are obtained as a function of consumption, by inverting the utility function. From (25) it follows that, given Jensen's inequality, the standard Euler equation for consumption will not hold. This is consistent with the fact that, in this world, all aggregate saving is done through the social planner. Like in the full information case we discuss above, as the social planner observes private savings, she can use aggregate wealth to maximize aggregate utility and adjust transfers to replicate any allocation achieved without private assets in a situation in which these assets are held by the individual consumers. As agents do not hold assets, there is no necessity for the standard Euler equation to hold. Equation (25), instead, holds as part of the dynamic incentive compatibility constraint.

In this world if consumers were left to their own devices, they would save too much compared to the social optimum corresponding to equation (25). This has important policy implications for optimal capital income taxation.24

Not many empirical papers have studied asymmetric information models of the type described here. An exception is Ligon (1998), one of the first papers that tries to discriminate between the self insurance (permanent income hypothesis) and the imperfect information partial insurance models. Ligon fits three different models by maximum likelihood using per-capita consumption data for households living in the three Southern Indian villages studied by Townsend (1994). Ligon points out that the only difference between the permanent income hypothesis Euler equation and the imperfect information inverse Euler equation (25) lies in the

---

24 An interesting result in this literature is the so called 'immiserization': because of the trade-off between incentives and insurance faced in each period by the planner, the efficient allocation equilibrium implies ‘an ever-increasing fraction of resources to an ever-diminishing fraction of society’s population’ (Atkeson and Lucas, 1992, see also Green, 1987, as well as Thomas and Worrall, 1990)
expected sign of the $h_0$ coefficient in the equation $E_i\left(\frac{C_i^{x+t,h}}{C_t,h}\right)^{h_0} = 1$: in the former case, this is minus the coefficient of relative risk aversion, in the latter case it is plus the coefficient of relative risk aversion. In the case of full insurance, this coefficient cannot be estimated, because there is no variability across households in consumption growth. The exercise can be carried out for a number of different instrument sets, and is consistent with permanent income behaviour in one village, and with imperfect information behaviour in the other two. Full information cannot be rejected in some cases.

In the standard asymmetric information model, it is often assumed that income or effort are not observable and contractable upon. At the same time, however, it is assumed that individual assets are fully observable. This assumption implies that, when considering the social planner problem that is often used to characterize the equilibrium in this class of models, one can dispose of individual assets altogether and assume that all assets are held by the social planner. Any equilibrium with private asset ownership can be replicated by an appropriate set of transfers. The situation becomes very different when not only individual income but also assets are not publicly observable. This problem has been analyzed in an important recent paper by Cole and Kocherlakota (2002). These authors show that in their model the constrained efficient allocation coincides with the one would occur when agents are allowed to save with a single asset paying a fixed interest rate.

When agents have the possibility of saving on their own, they will use this intertemporal margin. This implies two things. First, from a technical point of view, the Euler equation for consumption will have to hold (and indeed it becomes part of the incentive compatibility constraints). Second, agents, because of their ability to transfer resources over time, will have a strong preference for strategies that lead to high net present value transfers. Moreover, they will prefer, ceteris paribus, transfers that are front loaded. Finally, the relative preference for front loading will be stronger for agents with a low income realization than for agents with a high income realization. The implication of this will be that incentive compatible transfers cannot deliver less net present value to high income agents. But risk sharing would imply giving more net present value to low income agents. Therefore, incentive compatibility works exactly in the opposite direction than insurance. It turns out that in the constrained efficient allocation, all
agents will receive the same net present value that they will smooth using the hidden technology.

This result has been considered particularly important because it constitutes a micro foundation for a specific market structure (a single bond with a fixed interest rate) that has been widely used in the literature and that we have been discussing above. It would therefore seem that imperfect information about income and assets could provide a justification for a market structure where the only type of insurance agents can get is self-insurance, through savings. The Cole and Kocherlakota (2001) result, which is reminiscent of the results in Allen (1985), is an important one, even though some recent papers have claimed that it is not very robust. Abraham and Pavoni (2004), for example, have shown that in a model with pure moral hazard (rather than adverse selection of the type considered in Cole and Kocherlakota), one obtains their results only under very strong assumptions on the nature of the income process and of the moral hazard.

In a recent paper, Attanasio and Pavoni (2008), have shown that in a relatively general moral hazard model with hidden assets, the social planner can provide more insurance to the agents than in the bond economy mentioned above. In particular, the amount of risk sharing that can occur in equilibrium depends on the severity of the moral hazard problem. Attanasio and Pavoni (2008) write examples for which a closed form solution can be derived and in which the amount of risk sharing over and above that obtained by self insurance can be related to the degree of excess smoothness discussed in Section 3 and estimated on aggregate data by Campbell and Deaton (1989). In particular, these authors can interpret the degree of excess smoothness as reflecting the severity of the moral hazard problem. The larger the output loss involved with shirking, the easier it is to provide incentives and, therefore, insurance, and the larger is the degree of ‘excess smoothness’, i.e. the lesser is the response of consumption to innovations to permanent income. In their empirical work, Attanasio and Pavoni (2008) frame their test as a test of the intertemporal budget constraint, along the lines proposed by Hansen, Roberds and Sargent (1991). Risk sharing over and above the self-insurance provided by saving results in a violation of the intertemporal budget constraint with a single asset because it ignores the transfers connected with this insurance arrangement. Attanasio and Pavoni (2009) show evidence from UK micro data that is consistent with excess smoothness and interpret it within their theoretical framework.
Some important implications of incomplete and asymmetric information models are those about the optimal taxation of capital. In particular, some recent contributions have shown how imperfect information can provide a rationale for capital taxation, which contradicts the standard Ramsey result on optimal dynamic fiscal policy. Golosov, Kocherlakota and Tsyvinski (2003) show that in a model with unobserved and evolving skills it is indeed optimal to have a positive rate of taxation of capital income, although the Atkinson-Stiglitz result of uniform commodity taxation survives. The positive taxation of capital follows from the fact that the standard Euler equation does not hold, and instead the ‘inverse Euler equation’ does. The intuition behind this result is that, in these models, an increase in capital has two effects on welfare in the following period. On one hand, as in the standard model, it increases the resources available to the individual in the second period. On the other, it has a negative effect on incentives. In other words, it is optimal for the government to introduce a wedge between the interest rate paid to and the interest rate received by the consumer – by taxing interest rates the government reduces private savings.  

The exploration of models with asymmetric information and their implications for risk sharing and consumption behaviour is only beginning. There are very few empirical studies of this type of model, which constitute an exciting and important research agenda. Kocherlakota and Pistaferri (2009), whose paper we discuss in Section 5, is an example.

4.3.2. Imperfect enforceability of contracts

The other imperfection that the literature has explored as a possible reason for the lack of complete markets is the imperfect enforceability of contracts. There might be situations in which institutions that guarantee the execution of contracts are not developed enough and, as a consequence, individuals only enter contracts that are self enforceable. These models have been studied in a variety of contexts.

25 The paper by Golosov, Kocherlakota and Tsyvinski (2003) has been followed by other contributions that have looked at different aspects of this type of model. In particular, while this paper does not provide results on the decentralization of the constrained efficient allocations they study, Albanesi and Sleet (2006) and Kocherlakota (2006) study tax systems that could decentralize those allocations. More recently Fahri and Werning (2006) have considered the quantitative implications of models in which the inverse Euler equation holds.

26 Thomas and Worrall (1988) introduced some of the concepts used in this literature in the context of wage contracts. Kehoe and Levine (1993) introduced a very simple framework in which, in a deterministic world, simple idiosyncratic income fluctuations cannot be smoothed because of these enforceability problems. Kocherlakota (1996) analyzes a consumption problem, while Keoh and Levine (2001) extend their 1993 model to consider stochastic environments. Coate and Ravallion (1993) represents an early contribution which looked at the
Self-enforceable contracts can generate very rich dynamics and dynamic allocations that resemble some of the features seen in reality. In particular, one can generate interpersonal transfers of resources that are midway between insurance and loans. This type of allocations has been observed in several developing countries: Udry (1994) for instance, describes loan contracts in Northern Nigeria where the terms of the contract (interest rate and maturity) vary ex-post with the shocks received both by borrowers and lenders. Platteau, Murickan and Delbar (1985), instead, describing fisheries in Kerala, India, refer to quasi credit arrangements. In both cases, a distinctive features of the observed contracts and implied transfers, is that they are state contingent, like insurance contracts. On the other hand, they seem to have a memory, like debt contracts.

To understand how limited enforceability contracts may induce this type of transfers, it is useful to consider a simple example in which there are only two individuals, A and B. In such a situation, one can consider a social planner problem (like 16) and augment it with two participation constraints. These imply that each of the two consumers, in each history, prefers being in the contract to reneging and consuming its current income. With two consumers a given transfer will imply moving some resources from one consumer to the other. Clearly, only the participation constraint for one of the two individuals can be binding. This implies that only one of the two Kuhn tucker multipliers associated to the participation constraints is positive, while the other is zero. Thomas and Worrall (1988) and Ligon, Thomas and Worrall (2002) show that the behaviour of the two consumer economy can be summarized by a single state variable, which is the ratio of marginal utilities of the two consumers. Notice that under perfect risk sharing, this ratio is a constant and equal to the ratio of Pareto weights. In this context, however, this is not any longer true. Instead, when the participation constraint of one of the two consumers is binding, the ratio of marginal utilities will move. It is as if the social planner changes the consumption allocation relative to the one that would be observed under consequences of lack of enforceability for consumption allocations. The contracts considered in that paper, however, are not fully efficient, as they are restricted to be stationary. An important issue is the concept of equilibrium one uses. One wants to construct and characterize contracts that, in some sense, are self-enforceable. It is therefore necessary to establish what happens out of equilibrium, when somebody does not respect the terms of a contract. Most of the papers in this literature have resorted to the equilibrium concept proposed by Abreu, Pearce and Stacchetti (1990): when somebody deviates from the contract, the economy reverts to the worst sub-game perfect equilibrium, which turns out to be autarky. The punishment, on the basis of which some risk sharing is implemented, therefore consists in denying the benefits from future risk sharing. The amount of utility an individual can derive in the absence of risk sharing determines crucially the amount of risk sharing that can be sustained in equilibrium.
perfect risk sharing to guarantee that the constrained consumer is indifferent between staying in the contract and leaving. This implies rewarding the constrained consumer with a shift in promised utility. Effectively the Pareto weights that are an exogenous constant (which pick a specific competitive equilibrium) under perfect risk sharing, become endogenous and move to guarantee enforceability.

In such a situation it can happen that, if consumer A receives a sequence of consecutive positive shocks and is ‘constrained’ by the risk sharing agreement, the ratio of marginal utilities shifts progressively in her favour. It can therefore also happen that in some situations, consumer B, after receiving an income below her long-term average, actually transfers resources to consumer A who has experienced a positive shock. This type of behaviour makes the optimal contract, in this situation, resemble a debt contract: having borrowed from A, B is repaying some of her debt. However, the analogy only lasts until B is constrained by her participation constraint. When that happens, all past history is erased and the transfer of resources will be determined only by the necessity of keeping consumer B in the contract. When neither participation constraint binds, the efficient equilibrium will dictate that the ratio of marginal utility is kept constant and transfers of resources between the two consumers will guarantee that.

Clearly the nature of preferences and the properties of income processes will determine the amount of risk sharing that can be sustained in equilibrium. For instance, it can be proven that if the discount factor is high enough, perfect risk sharing will be sustainable, while if it is low enough, autarky is the only equilibrium. In general (but not always), an increase in the variance of income will lead to an increase in risk sharing, as it makes the value of autarky lower, while an increase in the persistence of idiosyncratic income will reduce risk sharing. This class of models has been applied widely: Alvarez and Jerman (2000) have studied the asset pricing implications of these models and stressed how the price of risk is determined by a subset of agents, while Kehoe and Perri (2002) have looked at the implications for international financial markets. In an interesting paper, Azariadis and Lambertini (2003) have introduced imperfectly...

---

27 It turns out that this type of situation can only occur if there are aggregate shocks.

28 Whether an increase in the variance of income causes a decrease or an increase in the amount of risk sharing depends on a variety of factors but, in particular, on how one increases the variance. If it is increased by expanding the range of income values it is possible to have a decrease in risk sharing. This is because the value of autarky for individuals at the right tail of the income distribution might go up if the discount factor is low enough. On the other hand, when the variance is increased by keeping the support of the income process unchanged and increasing the weights on the tails one gets an increase in the amount of risk sharing as the value of autarky declines.
enforceable contracts in an overlapping generations model. There, to get some risk sharing, punishment must imply a prohibition on saving, as well as participation in insurance markets.

There is not much empirical work on models with imperfectly enforceable contracts. One of the earliest contributions can be found in Foster and Rosenzweig (2001), who extend the dynamic limited commitment model to the case where consumers have altruistic preferences. They stress that altruism within extended families has an ambiguous effect on risk-sharing arrangements: there are greater utility gains from insurance, but scope for insurance is more limited if incomes of family members are highly correlated. Also, if altruism is very strong, the threat of autarky is no longer credible, and the mutual insurance scheme loses some of its appeal. Foster and Rosenzweig use transfer information data for the same three Indian villages analysed by Townsend, and show that imperfect commitment effects (generating history dependence) are generally important, but transfers are more responsive to shocks and less history dependent when income correlation is lower and altruism is moderate – as is the case of transfers to family members who live outside the village – in line with model predictions.

Ligon, Thomas and Worrall (2002) develop, solve and estimate a model with imperfect enforceability in a context where saving/borrowing is not allowed. They use the same Indian village data of Townsend (1994) and Ligon (1998), and carry out fully structural estimation of the three deep parameters that characterize the solution (subjective discount factor, relative risk aversion, and a state-independent punishment for reneging on the dynamic insurance arrangement), conditional upon the estimated income process. Two estimation procedures are carried out – one where the criterion function is in terms of the difference between observed and predicted individual log consumption, another one where the criterion function is instead the difference between changes of individual consumption shares over time. Even though estimated parameters take sensible values, they can either explain the distribution, or the dynamics, of consumption – not both.

Dubois, Jullien and Magnac (2008) add formal, incomplete contracts to a model of dynamic, limited commitment with storage. These contracts are meant to capture such arrangements as land-renting and share-cropping. They derive the efficient equilibrium allocation that is characterised by two equations; an Euler equation linking consumption growth to lagged consumption and current income because of the limited commitment insurance scheme (that introduces a borrowing restriction term in an otherwise standard equation), and an income
equation, where current income is affected by lagged consumption because of the formal contracts.

Dubois, Jullien and Magnac (2008) stress that the existence of formal contracts may either help or hinder informal transfers, as it affects both incentives and the possibilities without an agreement, a point also made by Attanasio and Rios-Rull (2000), and further developed in Albarran and Attanasio (2003). An important contribution of the paper by Dubois, Jullien and Magnac is the derivation of (non-linear) estimable equations for income and consumption – in their application to Pakistani village data they can thus avoid fully structural estimation, that is highly computer intensive (to the point that Ligon, Thomas and Worrall do not even report standard errors of their estimated deep parameters!). They show that the model is not rejected by the data, and that their estimated parameters imply that the probability of a self-enforcing contract binds is 10% if the relative risk aversion parameter is assumed to be equal to 1.5.29

Most if not all of the papers we have cited so far assume complete information. This implies the possibility of complete contingent markets: (self) enforceability is the only constraint that is imposed on the contracts available to an individual. These assumptions do not seem too strong for simple village economies, of the type studied by Udry (1998)30 and Platteau, Murickan and Delbar (1985). Whether they make these models relevant for developed economies is an interesting question. However, much has been learned from them. And interesting directions of research can be taken up. First, one can in principle try to introduce simultaneously information and enforceability problems. Atkeson (1991) considers one such model in the context of international financial markets. Phelan (1998) consider one sided lack of commitment and asymmetric information in a banking model. Alternatively, one can also introduce punishments that differ from the permanent exclusion from financial markets. Lustig and Van Nieuwerburgh (2005), for instance, have considered collateral constraints, in that individual consumers can only borrow against their housing wealth. This induces interesting effects of house prices on consumption and, more generally on asset pricing.

29 Other papers that have studied this class of models are Kehoe and Perri (2002) and Kehoe and Levine (2001)
30 Udry (1998), however, stresses the importance of private information in the region of Nigeria he studies.
5. Alternative models.

The standard model we have considered so far assumes that individuals solve a well specified optimization problem and the observed outcomes reflect, by and large, this type of behaviour. This approach might find it hard to explain some facts that have attracted attention in recent years, with the increased interest in the so-called behavioural economics. These facts generally refer to deviations from optimal behaviour. It has become fashionable to report anecdotes about apparently irrational or sub-optimal behaviour that ‘leaves money on the table’. In this section we discuss some of these puzzles and the evidence that generates them and try to put the issues in perspective.

Our general take is that a model of individual behaviour cannot fit the data perfectly and there will always be room for unexplained behaviour, which for lack of a better word, we define as ‘taste shocks’. Of course, if one were to find out that there are systematic deviations of observed behaviour from what is predicted by the model and that most observations need a ‘taste shock’, it would be an indication that the model is not a good representation of reality. The model would lose its predictive power and its usefulness. The issue is whether the model is able to capture some key features of individual behaviour and in particular the response to economic incentives. This is, in the end, what matters not only from a theoretical point of view, but also from a policy perspective.

We argue that in the case of life cycle consumption and saving decisions it is important to build models that are flexible enough to reflect the complexity of the environment and incentives that individuals face. It is also important to take into account the constraints (in terms of information as well as resources) that individuals are subject to. We will come back to these issues in the conclusions.

We start this Section by listing some of the puzzles that have been identified in the literature. We then discuss some extensions and modifications of the standard model that have been proposed in the literature.

5.1. Some puzzles

The literature has identified a number of facts that seem to be inconsistent with the standard versions of the life cycle model that we have discussed so far. Here we list some of those that have received most attention.
1. Inertial behaviour.

A series of recent papers (Madrian and Shea, 2001, Choi, Laibson, Madrian and Metrick, 2002, 2004, 2006) have documented that default options have important and surprising effects on the structure and level of saving of individual households. In particular, these papers have shown that if newly hired individuals are enrolled by default into a 401(k) retirement plan (rather than having to enrol), they are much more likely to participate, even though they have the possibility of opting out of the plan. In other words, the evidence seems to indicate that individuals with the same opportunity set make different choices depending on the default option they are (exogenously) assigned to. As the authors of these papers note, this fact contradicts the standard model where, in the absence of large adjustment costs, the default option should not matter. The authors of these papers propose a number of different explanations, ranging from the importance of ‘inertial behaviour’ and procrastination (possibly induced by the difficulty of the problem relative to individual ability to solve it) to the possibility that defaults are somehow perceived as a form of endorsement or advice, to ‘present bias’.

More recent papers, such as Carroll, Choi, Libson, Madrian and Metrick (2009) and Choi, Laibson, Madrian and Metrick (2009), show that inertial behaviour might extend to other related phenomena. The former paper, for instance, shows that, in addition to default, even forcing individuals to make an explicit choice (without giving them a default option) has an important effect on the decision to enrol in retirement plans. The latter, instead, show that the portfolio allocation of individuals with (exogenously) different allocations of the employer’s matched contributions is not systematically different, while it should be under a standard model.

2. Demand for commitment devices.

There is evidence that individuals are interested in devices that tie their hands in some relevant economic domain. For instance, Della Vigna and Malmendier (2006) discuss the choice of contracts offered by three health clubs, and taken up by their members: surprisingly large fractions of individuals choose to pay a flat monthly fee, but then rarely show up at the gym. For these individuals, the option of a 10-visit pass would work out to be much cheaper. An explanation is that consumers are willing to pay more for contracts that force them to do what is right for them in the long run, but is hard in the short run. Thaler and Benartzi (2004) report evidence on the first three implementations of a program whereby people commit in advance to allocating a portion of their future salary increases toward retirement savings. This programme has had a very high take up rate (almost 80%) and has led participants to save
much more than they used to before enrolment. Finally, Ashraf, Karlan and Yin (2006) report evidence from the sale of a commitment savings product for a Philippine bank that led to significant and lasting increases in savings by those customers who were offered it and purchased it. Thus there is evidence that consumers not only like to tie their hands in saving decisions, but that they save more as a result.31

3) Credit cards debt with low interest asset holdings.

Gross and Souleles (2002) were the first to point out that many households who borrow at high interest on credit cards have non-negligible investments in low-yield liquid assets. Laibson, Repetto and Tobacman (2003) report that among households with a head between ages 20-29 that are in the top wealth quartile, three-fourths did not repay their credit card bills in full. For households whose head in the thirties, over 80% of median wealth-holders had credit card debt. Even among the households with a head between ages 50-59 that were between the 50th and 75th wealth percentiles, 56% borrowed and paid interest on credit card debt in the past month. Laibson, Repetto and Tobacman (2009) conclude that “The typical American household accumulates wealth in the years leading up to retirement and simultaneously borrows on their credit cards”.

4) High saving rates in developing countries.

The life-cycle model explains cross-country differences in saving rates as follows: in high growth countries the young – who save – are life-time richer than the old – who dis-save. This explains why high growth countries save more. The very high saving rates in China have been explained by Franco Modigliani and Shi Larry Cao (2004), who point to the effects of the one-child policy on middle aged families in a country where most people cannot expect to receive a pension when they retire. The demographic imbalance implies that the traditional mechanism of intergenerational risk-sharing cannot be expected to provide adequate coverage for risks related to longevity and old age health problems. Private savings are then the only way to ensure an acceptable standard of living in old age. Recently, Chamon and Prasad (2007) have used micro data to analyze household saving rates by urban Chinese and found that the age profile of savings displays a U-shape. This is hard to explain in a standard life cycle model, where non-pension wealth should be decumulated in old age. The evidence that at least in some high-growth countries (China, Taiwan) the older generations save could be due to some form of habits (Paxson, 1996).

31 DellaVigna (2009) reviews a number of other papers in this area.
5) The equity premium puzzle (and low stock market participation).

The equity premium puzzle (Mehra and Prescott, 1985) has attracted much attention in the macro-finance literature. Given the historically high equity premium (the difference in expected return between stocks and bonds), asset markets equilibrium requires consumers to have very high risk aversion. This, in models with expected utility maximization and intertemporally separable preferences, in turn would imply a very low elasticity of intertemporal substitution, contrary to the empirical evidence and inconsistent with equilibrium conditions for the risk-free interest rate.

A number of possible solutions for this puzzle have been presented in the literature. One prominent hypothesis is the presence of habits in the utility function, as in Campbell and Cochrane (1999). We discuss habits, different ways in which they can be modelled and the evidence on their presence below. However, Bansal and Yaron (2004) claim that habits are not required to explain the key patterns of financial returns and consumption data, as long as one recognizes the existence of time-varying risk premia that generate consumption growth predictability. Bansal and Yaron, in particular, allow for the presence of a small long run predictable component in consumption growth and show that their model, with Epstein-Zin preferences, can explain several features of observed asset prices. More recently, some papers, including Lustig and Van Nieuwerburgh (2005), Chien, Cole and Lustig (2009), Lustig and Chien (2009) and Kocherlakota and Pistaferri (2009) have considered the asset pricing implications of models with asymmetric information and suggested that this class of models could be part of the solution to the equity premium puzzle.

An issue conceptually related to the equity premium puzzle is the limited participation into financial markets: in most countries relatively few households actively hold shares and equities. This has been labelled the stockholding puzzle by Haliassos and Bertaut (1995): given the relatively high returns on equities that have prevailed in many countries, households should invest at least some of their wealth in stocks. Even though this issue relates to the equity premium, it has not been addressed in the context of habits or imperfect information models, rather in models where consumers are affected by transaction costs or have access to limited financial information. 32

5.2. Modifying the basic model

5.2.1. Relaxing geometric discounting.

There is evidence, briefly reviewed earlier on, that individuals are interested in devices that tie their hands in saving decisions and that they save more as a result. The standard model of intertemporal decisions is at pains to explain this type of behaviour – where consumers apparently fear their lack of control and their inability to stick to their chosen optimal plan. Time inconsistent or temptation preferences have been proposed to rationalize some of the facts.

An elegant way to introduce time-inconsistent preferences is provided by the quasi hyperbolic discounting framework proposed by Laibson (1997) who developed ideas previously introduced by Strotz (1956) and Phelps and Pollak (1968). Consumers are assumed to maximize the expected value of the following life-time utility index:

\[ u(C_t) + \beta \sum_{s=1}^{T-t} \delta^s u(C_{t+s}) \]  

(26)

This implies that a different, lower discount factor is used to choose between this period and the next (the product of \( \beta \) and \( \delta \)) and between any two other periods (\( \delta \)), in agreement with experimental evidence provided by Thaler (1981) and Benzion, Rappoport and Yagil (1989). This discounting mechanism generates time inconsistent plans, with too little saving for retirement. Naïve consumers can do little about this, but sophisticated consumers recognize the problem and tie their hands, to prevent their current self from leaving their future selves in financial distress. This explains why consumers may choose to enter long-term saving commitment plans, such as 401(k)’s in the US (Choi, Laibson, Madrian and Metrick, 2006), or other committed saving products (Asharf, Katlan and Yin, 2006).

The quasi hyperbolic discounting model lends itself to estimation and testing, but requires solving for the consumption function numerically. Even though an Euler equation for this model has been derived, its empirical use is limited, because it involves the marginal propensity to consume out of wealth (Harris and Laibson, 2001). It also suffers from some potential difficulties related to the definition of the time period, that crucially affects the properties of the solution but the length of which is arbitrarily set by the researcher.
A more tractable and elegant specification of preferences that may be used to model quasi-rational impatience has been put forward by Gul and Pesendorfer (2001, 2004), who stress the importance of self-control problems leading to the postponement of saving. Their model can be characterized by a period-t utility function as follows (Bucciol, 2009):

\[ U_t = U(C_t) - \tau(U(CH_t) - U(C_t)) \]  

(27)

where \( CH \) denotes cash on hand (the sum of income and wealth), and \( \tau \) is a non-negative constant. The larger \( \tau \) is, the stronger the role played by temptation, inducing consumers to try and equalize consumption and cash in hand. (In this version of the model, consumption cannot exceed cash-on-hand).

Amador, Werning and Angeletos (2006) consider the issue of optimal trade-off between commitment and flexibility in a model where individuals expect to receive relevant information regarding tastes and thus they value the flexibility provided by larger choice sets, but also expect to suffer from temptation, with or without self-control, and value the commitment afforded by smaller choice sets. Their key finding is that imposing a minimum level of savings is always a feature of the solution. This has important implications for public policy – compulsory contributions to social security or other retirement saving schemes may be justified on welfare grounds even if actuarially fair annuities are available to consumers. The optimal size of these contributions will depend not only on the preference parameters discussed above (the elasticity of intertemporal substitution and the rate of time preference), but also on the impatience parameter in (27) or the self-control parameter in (28). In economies characterised by dynamic inefficiency, the presence of impatient or tempted consumers may help reduce the (excessive) amount of capital.

The hyperbolic discounting model finds its justification mostly in experiments. Frederick, Loewenstein, and O'Donoghue (2002) provide an overview of experiments that support that notion that individuals discount the near and the distant future at different rates. Rubinstein (2003) casts doubts on the interpretation given to the experimental evidence, by showing that differently designed experiments still suggest a marked preference for today versus tomorrow, but not in a way that is compatible with hyperbolic discounting. Recently, Fernandez-Villaverde and Mukherji (2008) claim that the role played by uncertainty biases results against exponential utility. To counter this criticism, Ameriks, Caplin, Lehay and Tyler (2007) devise a set of questions aimed at eliciting self-control problems directly and claim most people are affected by such problems (less so as they age). Fernandez-Villaverde and Mukherji (2008) devise a
different experiment, where sophisticated hyperbolic discounters should take a commitment device, while exponential discounters and naïve hyperbolic discounters should not. They report that only 13% choose the commitment device in their experiment.

In recent years an increasing body of evidence comes also from estimation. Laibson, Repetto, and Tobacman (2009) follow a fully structural approach and show that the sophisticated hyperbolic discounting model can reconcile credit-card debt with illiquid asset holdings over the life-cycle. Bucciol (2009) follows a similar approach, but estimates the temptation model instead. To identify the parameters he uses liquid and quasi-liquid (retirement) wealth holdings at different ages as target moments. He finds evidence of a small, but significantly positive degree of temptation – when temptation is taken into account risk aversion is found to be less (and statistically different from) one. 33

5.2.2. Relaxing intertemporal separability.

The standard model presented in section 2 assumes preferences to be additive over time and over states of nature: this implies that risk aversion and intertemporal substitutions are functionally related. In the special case of the isoelastic function, the relative risk aversion coefficient is the reciprocal of the elasticity of intertemporal substitution. The assumption of intertemporal separability, however, might be too strong as it cannot capture phenomena such as habits and durability.

The simplest way to introduce habits/durability of consumption is to write the utility function as follows:

\[ u(x_t - \theta' x_{t-1}; z_t) \]

where \( x \) is a vector of goods or services and \( z \) is any other variable that affects marginal utility (demographics, leisure, other goods that are not explicitly modelled). The \( \theta \) parameters are positive for goods that provide services across periods (durability), negative for goods that are addictive (habit formation) or zero for goods that are fully non-durable, non-habit forming (Hayashi, 1985).

Eichenbaum and Hansen (1990) and Ogaki and Reinhard (1998) test and reject the separability of durables and non durables within the context of an Euler equation estimated on aggregate data. Alessie, Devereux and Weber (1997) and Padula (1999) tests and rejects separability.

33 Mastrobuoni and Weinberg (2009) report that Social Security benefit recipients without savings (about a fourth of the sample) consume 25 percent fewer calories the week before they receive checks relative to the week afterwards. They show that their findings are consistent with hyperbolic discounting.
between non-durable consumption and the stock of cars using micro data from the FES and CEX, respectively.\(^{34}\)

Habits have attracted much attention in the macro-finance literature. In the presence of habits, the functional restriction between intertemporal substitution and risk aversion is relaxed. Campbell and Cochrane (1999) make the distinction between the overall curvature ($\gamma$ in the isoelastic case), that is relevant for intertemporal allocations of consumption, and the local curvature of the utility function, ($\gamma$ over the surplus consumption ratio, that is the share of consumption net of habits over consumption), that is instead relevant for portfolio decisions.

Habits can take various forms: today’s marginal utility may depend on the consumer’s own past consumption level (internal habits) or the past consumption level of other consumers (external habits). Campbell and Cochrane consider the case of external habits, where consumers are influenced by other households’ lagged consumption, not their own, and show that their model can solve the equity premium puzzle for plausible parameter values. The external habits model seems to work better than the internal habits model on aggregate data, even though Chen and Ludvigson (2009) challenge this conclusion.

Empirical macro-evidence on the presence of habits is quite mixed, and this may be due to the very nature of aggregate consumption data, as stressed in Dynan (2000). The serial correlation of aggregate consumption growth is affected by time aggregation (Heaton, 1993), aggregation over consumers, and by data construction methods (particularly for the services from durable goods). For this reason micro data seem preferable.

The Euler equations corresponding to (28) involve $x$ at four different periods of time, and their estimation typically requires panel data. High quality consumption panel data are rare, and this has limited the scope for empirical analysis. Meghir and Weber (1996) have used CEX quarterly data on food, transport and services (and a more flexible specification of intertemporal non-separabilities than is implied by equation 28), and found no evidence of either durability or habits once leisure, stock of durables and cars as well as other conditioning variables are taken into consideration\(^{35}\).

\(^{34}\) A special feature of durable goods is that they might be subject to adjustment costs. The case of convex adjustment costs is a relatively simple extension, and captures well repairs and maintenance activity (see Bernanke, 1985). To model replacement decision, non-convex adjustment costs are more plausible, as they lead to infrequent adjustment, and explain why durable goods are not replaced all the time – the seminal paper is Grossman and Laroque (1990) – applications are Eberly (1994), Attanasio (2000) and Bertola, Guiso and Pistaferri (2005).

\(^{35}\) Indeed, Flavin and Nakagawa (2008) argue that the presence of non-separable, illiquid durable goods such as housing in a standard utility function explains the smoothness of aggregate non-durable consumption the same
Similarly negative evidence on habits has been reported by Dynan (2000), using PSID annual food at home data. On the other hand, Carrasco, Labeaga and López-Salido (2005) use Spanish panel data that follow households over eight consecutive quarters and find evidence for habits, once they control for fixed effects.

Habits are more likely to explain the high saving in developing countries puzzle if they persist over a long period. Angelini (2009) has worked out the analytical solution of the dynamic optimization problem when preferences are CARA and there are habits in the utility function. An interesting feature of the solution is the interplay of habits and the precautionary saving motive – another is the dependence of beginning of life consumption on “inherited habits” – a feature that could be exploited in empirical work in data sets that contain information on the standard of living enjoyed early in life, before leaving the parental home.

The few studies that have used micro data on non-durable consumption items to investigate the issue find little or no evidence of habits, at least once preferences capture the presence of non-separabilities between goods and leisure.

5.2.3. Financial literacy and information.

A standard assumption in the life cycle literature, made mainly for analytical and empirical convenience, is that of rational expectations. This assumption states that individuals know the stochastic environment in which they live, have at least as much information as the econometrician in making their consumption and saving decision and use it optimally. Recently much evidence has been gathered that sheds important doubts on this assumption. A number of papers (see, for instance, Lusardi and Mitchell, 2007, 2009, Lusardi and Tufano, 2009) have used explicit and quantitative measures of financial literacy and related them to individual financial decisions. Lusardi and Mitchell (2007, 2009) show that more ‘financially literate’ individuals are more ‘retirement ready’, while Lusardi and Tufano (2009) show that more ‘financially literate’ youths are less likely to hold unsustainable debt. It has also been shown that stock holdings are much less common among the less financially literate.

Perhaps it is not surprising that financial sophistication affects individual behaviour. It could also be that financial literacy is correlated with other individual attributes (such as total human way as the external habits model. This could reconcile the failure to find micro-evidence on habits with the success of the external habits model to fit the aggregate consumption and financial returns data.
capital) that are linked to the amount of resources an individual controls (a fact recognized by some of the papers in the literature). However, the evidence that has been gathered so far is reasonably convincing and should be taken seriously.

6. Conclusions

The aim of this paper was to survey the theoretical and empirical literature on the life cycle model to draw the implications that plausible versions of the model have for public policy and, in particular, for policies that influence the intertemporal allocation of resources. Rather than summarizing what we have been discussing above, we conclude this paper by taking a stand on what we think are profitable directions for future research.

One possible reading of the empirical literature on the life cycle model is that it is possible to construct rich versions of the model that are not inconsistent with available micro data, especially for households headed by prime aged individuals. Much of this evidence comes from the estimation of Euler equations. Euler equations are remarkably useful because they let researchers estimate important preference parameters in a relatively robust way, allowing for - but without the need to explicitly model - important phenomena such as labour supply, housing, durables and so on.

However, to conduct a useful policy debate it is necessary to be able to say something about the level of consumption. A reduced form approach that exploits key theoretical insights can shed light on some issues – for instance on the nature of particular business cycle episodes or whether consumers perceive specific shocks to be permanent or temporary. A structural form approach is more generally informative, but requires, except for special cases, numerical methods and simulations. Moreover, it requires specifying completely the environment in which economic agents operate, including their perceptions and information sets, institutional factors such as pensions and intertemporal trades available to them. The necessity to provide so much detail makes this approach inherently not robust. This is not to deny its usefulness, but to make it clear that the general validity of results obtained using simulated life cycle models is not to be taken for granted.

Many of the features to be included in the model to make it realistic involve important non-convexities that make the optimization problem difficult to solve numerically. Much progress has been made since the first numerical simulations of life cycle models with uncertainty were developed by Deaton (1991). Current models are able to consider in very sophisticated fashion
housing choices, labour supply, liquidity constraints and a number of other factors. Much work, however, remains to be done to develop these models. Moreover, while some of the parameters of these models can be estimated by Euler equations, many of them cannot and one has to obtain sensible estimates of crucial parameters from alternative sources.

There is still work to be done in term of understanding intertemporal preferences. The work on Euler equation has made it clear that one needs to take into account the evolution of individual needs and the non-separability between consumption and labour supply. Another aspect that has received less attention, but could turn out to be important, is the role played by durables and semi-durable commodities. There is some evidence of non separabilities that could be important in assessing individual responses to different shocks and innovations.

The Euler equation approach is useful because it can allow for the presence of these non-separabilities even when some of the choice variables are affected by non-convexities and other imperfections. The Euler equation provides equilibrium conditions that the policy functions that determine consumption and other choice variables have to satisfy and that can be used to estimate structural parameters even if we cannot characterize these policy functions explicitly. However, to understand the intertemporal allocation of resources, how individuals smooth shocks, how they react to policy innovations these policy functions become essential. Studies that do this in a systematic fashion are few and far in betweens. In our opinion, an important paper is Browning and Crossley (2009) that looks at how individuals use the timing of the purchase of durables to smooth out specific transitory shocks. Much more work is necessary in this direction. The recent recession and some of the policy measures taken in the UK can supply important examples of questions to which policy makers would like to have answers that economists are still unable to provide. For instance, if one lowers temporarily the rate of indirect taxation, what is the effect on consumption and, in particular, on durable purchases? And how does the answer change when the decrease happens in response to an increase in the level of uncertainty in the economy?

Another important potential use of this class of models is to study aggregate consumption and saving and, possibly, to construct realistic equilibrium models. The work on this is still in its infancy and faces some severe problems. The life cycle model is an intrinsically dynamic model in which choices depend on future variables. The equilibrium values of these variables depend, in turn, on the behaviour of all consumers in the economy. It is therefore difficult to establish what determines equilibrium values. Other equilibrium phenomena that are important and interesting are the determination of the type of assets that are available to individuals both to
smooth income shocks and to finance investments (such as human capital accumulation) when
information problems (adverse selection and moral hazard) are important.
The explicit modelling of imperfections and frictions that cause markets to be incomplete is
highly promising and potentially very useful in characterizing the implications that a structure
such as the life cycle model has for policy. The recent exciting development of the new
dynamic public economics is a good example of that.

In section 5 we have mentioned a number of directions in which some strong assumptions
routinely made in the literature on the life cycle model can be relaxed. The list contained in
Section 5 is not exhaustive. However, we chose those topics that we think are more promising
in terms of future research. The analysis of alternative preference structures, such as the
analysis of temptation in Gul and Pesendorfer (2004), is very important and so is the
consideration of habit formation.
As for the model that relaxes geometric discounting, systematic empirical studies are still rare.
The situation is slightly better in terms of habits, although the evidence based on micro data
still comes from a handful of studies. The biggest limitation so far comes from the fact that
many of these studies only consider habits with a very short duration, not necessarily because
this is the most appealing model of habits but because of data limitations. The study with
longer lasting habit stocks should be a high priority. Research on habits is likely to be important
both to assess the plausibility of the claims that habits can be helpful in explaining some of the
puzzles in finance and because they could explain the evolution of saving rates in fast-growing
developing countries, such as China.
Another area that deserves mention is the analysis of the role of financial literacy. It is clear
that in some aspects the standard model imposes very strong assumptions on the ability of
agents to solve intertemporal optimization problem and on the information they have at their
disposal. One possible alternative is to collect data on the information individuals agents have
and on the information they act upon when making intertemporal choices (both saving and
investment choices). If we cannot assume that agents are fully rational and have all the
necessary information, proper modelling and empirical work requires measures of beliefs and
expectations. Progress has been made in this direction (especially in measuring expectations).
But much more work is necessary. In our opinion, measuring financial literacy (and its
determinants) is an important direction of research. The same applies to individual beliefs,
attitudes, and preferences. In the same way in which the development of survey methods has
allowed in recent years a much more precise measurement of household financial wealth and (more recently) subjective expectations, we need to develop similar methods for the measurement of these other objects that are obviously key determinants of individual choices. Integrating these measures within rigorous but flexible structural models can yield high returns in terms of academic research and information useful for the design of effective policies.
References


Chien Yi-Li, Harold L. Cole and Hanno Lustig, 2009."Is the Volatility of the Market Price of Risk due to Intermittent Portfolio Re-balancing?," NBER Working Papers No. 15382


Fernández-Villaverde, Jesús and Arijit Mukherji, 2006. “Can We Really Observe Hyperbolic Discounting?”, WP


81


