Four studies of economic behavior: integrating revealed and stated preferences data

ISBN 978 90 5335 274 8
Printed by Ridderprint B.V., Ridderkerk

Four studies of economic behavior: integrating revealed and stated preferences data

Vier studies van economisch gedrag: een integratie van data over ”wat mensen doen” en ”wat mensen zeggen”
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof.dr. J.C. Stoof, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op woensdag 14 april 2010 des middags te 4.15 uur

door

Joppe Jaitze de Ree

geboren op 29 oktober 1975
te Groningen

Promotor: Prof.dr. R.J.M. Alessie

Contents

Acknowledgements

I started writing a Ph.D. thesis for two broad reasons. First, I was interested in understanding the processes leading up to (some) real world phenomena. Second, I was unable to translate the econometric methodology I learned before to achieve this ultimate goal. Currently, I am still interested in understanding the processes leading up to (some) real world phenomena, but now I have a pretty good idea of how to apply statistical and mathematical principles to study them. In my view therefore, my time as a Ph.D. student has been an absolute success.

There are many people who have played an important role in the realization of this thesis and in my development as a researcher more generally. A few people however need to be mentioned specifically. First and foremost I would like to thank Rob Alessie who has been my advisor on this project. From many discussions with fellow Ph.D. students I learned that the amount of energy that Rob invested in me is far from standard. This has been tremendously important. (In weekends and evenings, with pizzas or mixed grill from one of the many restaurants in Rob's neighborhood we discussed many issues usually accompanied with a beer or two).

Before I became a Ph.D. student at Utrecht University I was an intern at the Dutch Central Bank. Here, I once more realized that econometric (i.e., statistical) knowledge *alone* is barely adequate to study many complex real world phenomena. At the Central Bank I got involved with studying the following question: what are the effects of wealth holdings on household consumption? Surely, there might have been a few quick-and-dirty answers[1] to this question, but I felt that the truth was much more complex. It occurred to me that I had *no* prior idea of the possible mechanisms that *could* make people consume more, or less, out of wealth. It was there that I developed an interest in the economic theory of individual decisionmaking, based on

---

[1] The (partial) correlation between wealth holdings and household consumption would not be sufficiently informative to fully answer this question.

the principle of utility maximization.

This is where Rob came in. With enthusiasm he introduced me to the richness and usefulness of this methodology. It is this literature that offers a way to look at the world *and* a way to use econometric methodology in practice that I was searching for. Moreover, Rob did not get nervous when it took me approximately two years (!) to broadly understand the depth of the "life cycle model". Instead, Rob was continuously available for discussing this model, and many more of the basic principles of economics and econometrics. Furthermore, I thank Rob for the scientific freedom he provided me with, and for the trust in my abilities that he thereby revealed. Two examples: first, Rob supported my wish to work around the clock on a paper about civil war for more than a year, even though this could –and it eventually did– delay the delivery of this Ph.D. thesis. Second, he trusted in my nonstandard ideas about using subjective data in economics and supported my intentions to work them out, even though these projects were very risky. One easily gets used to having a great advisor, but nevertheless, I greatly appreciate it.

The second person that I want to thank is Maarten Bosker. Maarten and I grew up together in the world of economic science and we seem to agree on a lot of things. One might wonder whether this is *because* we grew up together or whether it is just coincidence. Whatever it is, it is undoubtedly true that my "scientific world" would have been a different place if Maarten would not have been part of the first shipload of Utrecht AiO's. Maarten, I have benefited from your down-to-earth mentality and free-spirited outlook on research. One of the things that I learned from you (and that came in extremely handy for finishing this thesis) is to stop worrying at some point and start writing stuff down! Hopefully we can prolong our partnership for another while. At least we do not have to worry about great (!) ideas for research, I mean, we haven't even had time to derive the formula for the spatial Nickell bias yet. It seems that the "dubious research projects" that we were thinking about a few years ago are actually leading to something, slowly but surely.

As I said before, there exists a long list of people that at various stages have contributed to the realization of this thesis. These people certainly involve my (former) colleagues in Utrecht and Berlin, my great group of friends in Amsterdam (and elsewhere), Adama, and my family, most notably my father and my mother who have always supported me in pursuing my own ideas. Finally, I would like to thank the members of the reading committee for helpful comments and your approval: Arie Kapteyn, Arthur van Soest, Harry Garretsen, Menno Pradhan and Adriaan Kalwij.

Joppe de Ree
Amsterdam
March 12, 2010

Chapter 1

Introduction

Applied microeconometrics is the science of inferring individual preferences from survey data of those individuals. Having knowledge of those preferences is important as it can be utilized to anticipate responses to government policy, or to help understanding why some people are poor and others are rich. The standard methodology of inferring preferences from survey data relies on the theory of revealed preferences (Samuelson 1938), where individuals reveal their preferences by undertaking actions in an environment of scarcity. If a child is offered a choice between an apple or a chocolate bar, it *reveals* its preference by choosing one, instead of the other. In the microeconometric practice, data on observed actions (e.g., consumption, labor supply or savings) are combined with data on choice sets (e.g., functions of prices and total outlay) to infer individual preferences.

Fairly recently, some attention of microeconometricians has shifted away from the revealed preference approach to another type: the stated preferences approach. It is intuitive that individual preferences cannot only be inferred from observed *behavioral* action, but also from *verbal* action: a researcher may simply ask individuals why they do, or appreciate certain things. This possibility sounds obvious and it is remarkable that economists have remained sceptical about this approach for a long time (there are exceptions such as Van Praag (1968)). There existed a widespread belief among economists that self-reports are merely cheap talk. Data on actual decisions were considered much more ”real” and therefore much more informative. There may be some truth in this argument. Yet, apart from the possibility that self-reports are just measuring noise, it is a potentially useful source of *additional* information. Especially, because stated preferences data are, to some extent, a richer source of information than data on observed actions.

So far, both strands of the microeconometric literature –the revealed preference and the stated preference approach– have been largely existing side by side. This fact may be partly explained by the absence of a coherent way of jointly interpreting the various types of self reported data and the data on observed actions.[1] Several different measures of satisfaction are usually bluntly referred to as utility (income or

---

[1] I typically refer to self reports on wellbeing or satisfaction. Self reports on income expectations are perhaps more straightforward to interpret within the standard microeconomic paradigm.

life satisfaction most prominently). This does not constitute a problem per se, if one could just take this particular concept of utility to be different than the one that is usually adopted in the revealed preference practice. I would argue however that this coexistence is undesirable for two broad reasons: *1.* if both ways of measurement say something about *the same* preferences one can increase efficiency by using both sources of information. *2.* If both ways of measurement have something to say about *different* preferences, they could both contribute to inferring a much more complete picture of preferences. In this line of reasoning Beshears et al. (2008) argue: "successful models of human decision-making should be able to explain both behavior and self-reports".

In general one may regard stated preferences data as an additional source of information to solve problems of identification that can be encountered in revealed preference approaches. Identification problems in general can be "solved" by imposing additional (identifying) assumptions or by using additional data. Manski (2004) for example, elaborates on the typically maintained assumption of rational expectations in the revealed preferences literature and argues that measurement of subjective expectations can be used to "relax or validate assumptions about expectations". Jappelli and Pistaferri (2000) use subjective expectations of income (i.e., stated preferences) as an instrument of income growth, in order to test the validity of the life cycle-permanent income model. (A basic version of) the life cycle model predicts that consumption growth is independent of anticipated changes in income. It seems clear that testing this prediction is greatly facilitated with self reports on anticipated income changes. Anticipated income, as a concept, cannot be readily inferred from observed actions. Without this kind of data, assumptions must be made to *model* anticipated changes in income.

For many additional purposes, a combination of revealed preferences data and stated preferences data can be useful. An important part of this thesis contributes to ongoing research that aims to combine both sources of information in a coherent way. Whereas a combined (economic) theory that explains both stated preferences and observed actions is currently underdeveloped, I discuss two possible directions in which stated preferences data are useful. Both directions feature in chapter 2 and chapter 3 of this thesis and both aim to solve identification problems that are encountered in revealed preferences practices.

First, the revealed preferences methodology has no power in situations where no actions are taken. In these situations, individuals have no chance to reveal their preferences by doing. Stated preferences data instead can be informative where the intended actions of individuals are obstructed by institutional factors such as liquidity constraints. Liquidity constrained individuals cannot borrow against future labor income to smooth consumption. Consequently, they do not reveal their preferences for smoothing consumption. This information problem may be resolved by asking a set of well-chosen questions to assess their willingness to borrow in relation to the restrictive policies of banks. Similarly, stated preferences are informative when preferences do not relate to behavior more generally. Life- or other measures of satisfaction are important elements of individual preferences, yet they do not necessarily relate to action. In chapter 2 I use stated preferences data to assess household wellbeing under different demographic regimes.

Second, stated preferences data can be useful for testing theoretical principles. The revealed preferences methodology *in practice* often relies on the validity of *some specific* utility function (that satisfies the axioms of rational choice). The data is subsequently interpreted within the context of that particular utility function. Statistical tests on the validity of the model, consequently, are tests on the validity of a specific parametrization of a broader theoretical idea. This fact very much limits the possibility for rejecting theoretical ideas and leaves (too) much room for debate: are the functional form assumptions inadequate or is it the theoretical idea that is inadequate?

It is often seen in practice that rejections of "simple" models are succeeded by more general systems of preferences that are able to withstand the statistical tests. Tests on *extended* models however, may suffer from low power, because many (preference) parameters need to be estimated simultaneously with the test statistic. Testing the validity of highly flexible models is therefore burdensome using revealed preferences data. Here also, stated preferences may help. Ingeniously designed questionnaires may directly question individuals about important elements of a theory. The advantage is twofold. First, one does not rely on the simultaneous estimation of structural parameters and test statistics. Second, the set of alternative hypothesis that is associated with a rejection can be greatly limited (i.e., functional form issues are no longer a cause of a rejection). In chapter 3 of this thesis I utilize stated preferences data to directly test the validity of the life cycle model, without relying on arbitrary assumptions on functional form.[2]

The chapters 2 and 3 present my progress in the field of combining the two sources of information. Chapter 2 estimates equivalence scales and chapter 3 uses stated preferences data in a statistical test of the life-cycle model. The other two

---

[2] Recent developments in the microeconomic literature discuss situations where observed actions do not match individual normative preferences. Normative preferences are the actual interests of people (Beshears et al. 2008). Individuals might want to stop smoking, but they effectively do not because they procrastinate. The fact that individuals keep smoking does not necessarily mean that they prefer smoking to quitting in a broader sense. This "time inconsistent" behavior predicts a discrepancy between behavior and individual interests (O'Donoghue and Rabin 1999). In situations where individuals are boundedly rational, observed actions do not always match individual preferences (complexity, hyperbolic discounting, myopia, etc.). Designing subjective questions to address this discrepancy is an interesting avenue for future research.

substantive chapters 4 and 5 of this thesis do not employ stated preferences data and contribute to the revealed preference literature. Chapter 4 describes patterns in important economic and demographic variables based on a survey among Dutch households and chapter 5 studies when and how much Dutch households choose to work, save and consume.[3]

In chapter 2 I present a study that estimates household equivalence scales using a sample of Indonesian households.[4] Particularly in the development practice, policymakers are interested in making welfare comparisons between households of different demographic composition. Due to the existence of economies of scale within a household however, such comparisons are not straightforward. Individuals within households share things like heating or a television (i.e., nonrival goods), but food and clothes for example are individual specific (i.e., rival goods). A consequence of this is that both household income and household income per capita are inconsistent measures of welfare. Equivalence scales are defined as household specific index numbers that can be used to scale *nominal* household income and transform them into appropriate measures of welfare.[5] These equivalence scales, therefore, take account of the economies of scale.

This chapter exemplifies the power of using demand data (i.e., revealed preferences data) and subjective evaluations of household welfare (i.e., stated preferences data) simultaneously fur the ultimate goal of this study: to estimate the price and utility (i.e., welfare) dependence of equivalence scales. Pollak and Wales (1979) show already that equivalence scales are not identified from observed patterns of demand. One needs, for example, subjective reports on wellbeing under different demographic circumstances as an additional source of information. However, where subjective data theoretically *identifies* all the important parameters of the cost function, and hence the equivalence scales, it is insufficiently *powerful* to do this with reasonable precision: on the basis of subjective data alone we would for example, spuriously conclude that equivalence scales are independent of prices and reference utility. Using revealed preference data –demands over food and nonfood items in this case– as an additional source of variation we conclude the exact opposite: we find overwhelming evidence for the statistical importance of the price and utility dependence of equivalence scales.

Whereas we find evidence for the *statistical* importance of the price and utility dependence of equivalence scales, the *economic* importance of these dependencies is arguably small. The magnitudes of the equivalence scales we estimate are "reasonable", in the sense that they are only a little larger than the well-known modified

---

[3] The chronological order in which the four substantive chapters of this thesis were written is: chapter 5, chapter 2, chapter 4 and chapter 3.

[4] This chapter is joint work with Rob Alessie and Menno Pradhan.

[5] "Appropriate" measures of welfare, in this context, can be used to make welfare comparisons between households of different demographic composition.

OECD scales (De Vos and Zaidi 1997).

In chapter 3 I test the validity of the life cycle (or expected utility) model, or more precisely, the validity of the Euler equation of food consumption which is implied by the life cycle model. For this I use the same household survey panel data set of Indonesian households that was used in chapter 2. This chapter employs the potential of stated preferences data in a different way. A (fairly general) life cycle model predicts, among other things, that marginal utility of food consumption remains constant over time, in expected discounted terms. That is, using clearer, everyday terminology: in expectation households are expected to "tailor their consumption patterns to their needs at different ages (Deaton 2005)" by saving when income is high and by borrowing when income is low.

In the literature, empirical tests on the Euler equation are typically operationalized as follows. First, parametric assumptions on the marginal utility function are imposed. Second, the parameters of this parameterization and the test statistics are estimated simultaneously. The simultaneous estimation of preference parameters and test statistics posits a problem. When baseline models have been rejected, newer, more elaborate versions are introduced that withstand statistical testing. These more flexible models have more ways of being true, and hence, less ways of being rejected. For example, Blundell et al. (1994) find that excess sensitivity of consumption to anticipated changes in income (i.e., a rejection of the baseline life cycle model) can be explained by nonseparabilities with demographic variables (i.e., failure to reject the demographically extended life cycle model). If the evolution of family size, goes hand in hand with the evolution income, empirical researchers will have a difficult time of rejecting the *demographically extended* life cycle model, even if it is an invalid representation of preferences.

In this chapter I use a direct, self-reported, proxy of the marginal rate of substitution of food consumption in two consecutive periods to test the validity of *the* life cycle model. Indonesian households responded to a question about *a change in the adequacy of food consumption over the past year* on a five point scale. The key assumption of this research is that the answer to this question is a –monotonic, but noncontinuous– transformation of the marginal rate of substitution of food consumption in two consecutive years. The advantage of using the proxy in an empirical test of the life cycle model is that: *1.* a rejection is associated with a smaller set of alternatives, as functional form issues such as curvature, nonseparabilities cannot explain a rejection. *2.* the approach does not suffer from the fact that test statistics and structural parameters are estimated simultaneously, hence, it can discriminate between theories that both produce very similar behavior. I find a borderline rejection of the predictions of the life cycle model with individual specific discount factors. I do not however, find evidence for the importance of binding borrowing constraints.

The fourth and the fifth chapter use data from the expenditure survey of the Netherlands.[6] This data set is a time series of cross sections that spans from 1978 to 2000. Both chapters are different from chapter 2 and chapter 3 and do not use self reported or stated preferences data. Chapter 4 reports life cycle profiles of some key economic and demographic variables, after controlling for time and birth cohort effects.[7] We look at household income, durable and nondurable consumption (at the household level and per adult equivalent), labor supply variables, number of children and number of adults within the household. This chapter describes the data without superimposing a preference structure on the data and therefore allows for an ad hoc and straightforward interpretation.

At the same time, the data can be loosely linked to life cycle theory: it is unlikely that a expected utility/life cycle model will generate a hump in consumption over the life cycle, after controlling for household demographics. This chapter documents a hump shape in both measures of consumption. After controlling for household size however, the hump shape largely disappears. Nevertheless, we find that nondurable consumption increases with income in the first phase of the life cycle, yet, it does not drop after retirement. Durable consumption shows an entirely different pattern. We find that durable consumption on average is constant until retirement, after which it drops sharply. Whereas we do not find a hump shape in both measures of consumption per adult equivalent, the findings still seem largely inconsistent with some basic predictions of the life cycle model. The findings however, facilitate discussions about where and when households seem to smooth consumption and where and when they seem to do something else. Elasticities of intertemporal substitution, or other specific economic concepts cannot be estimated using this approach. One could however, question the usefulness of these concepts if the life cycle model is an invalid representation of preferences.

Chapter 5 specifies and estimates the parameters of a structural dynamic model of female labor supply and consumption. The model is estimated in two steps (using the principle of two-stage budgeting). First we estimate a (life cycle consistent) within period model of female labor supply, and second we estimate a Euler equation of nondurable consumption using a long T synthetic panel of cohort averages. Using the estimated parameters we estimate static wage elasticities of female labor supply of around 1. If we take the intertemporal reallocation of resources into account that is typical for life cycle models, we estimate elasticities of around 1.7.

Additionally, we develop a theoretical novelty on how home production can be allowed for when data on home production and/or time-use *is not* available. We show that if the domestic production function exhibits constant returns to scale and when

---

[6] Both chapters are joint work with Rob Alessie.

[7] A restriction on the time dummies is imposed for identification.

the two factors of production (male and female time) are perfect substitutes, standard models of labor supply may still be consistently estimated, and interpreted. Under these assumptions one should *re*interpret the standard dichotomous tradeoff between leisure and consumption as a tradeoff between nonmarket time[8] and consumption.

Note that the research of chapter 5 is clearly an exponent of "revealed preferences theory in practice", with the disadvantage that the estimated elasticities are only valid if the model is correctly specified. The strong (identifying) assumptions we impose at the outset (e.g., separability between durable and nondurable consumption, the validity of the life cycle model) perhaps compromises the straightforward way of interpreting such parameters.

---

[8] Nonmarket time is defined as time spent leisuring or producing domestically produced goods, like child care.

Chapter 2

The price and welfare dependence of equivalence scales: evidence from Indonesia

Abstract

In this chapter we are estimating equivalence scales. We place particular emphasis on the price, and reference utility dependence of these scales. In order to estimate the necessary parameters, we analyze demographic effects on expenditures and self-rated consumption adequacy simultaneously within the context of an Almost Ideal Demand System. We find equivalence scales that are of reasonable magnitude (they are little larger than the well-known modified OECD scales). Moreover, we find that equivalence scales increase in the food to non-food price ratio for low welfare households, where we do not find price dependence for high utility households. More surprisingly, we find that equivalence scales increase in reference utility on meaningful price domains. Whereas both price and reference utility effects are highly important in statistical terms, their economic importance is arguably small. For practical purposes, "rule-of-thumb" equivalence scales, that are just functions of demographics, may suffice.

2.1 Introduction

To derive measures of inequality or of poverty, policy makers and social scientists make welfare comparisons across households of different sizes and compositions. For many purposes however, the two most obvious proxies for household welfare – household income and household income per capita– are too crude to be useful. Both measures are typically inconsistent in the sense that both measures –in general– do not uniquely reflect a *standard of living.* Ideally, applied welfare analysis is based of a utilitarian concept of welfare that takes into account the differences in household needs due to size and composition.

The utilitarian welfare concept can be operationalized by scaling (nominal) household income with appropriate household specific index numbers known as equivalence scales. In general, equivalence scales are functions of prices, demographics and utility itself, and take into account that for reaching a certain utility level a family of four needs more than a family of two, but generally not twice as much. Household incomes that are scaled with equivalence scales are subsequently referred to as equivalent incomes.

Equivalent incomes are consistent measures of welfare in the sense that ordering households on the basis of equivalent income *or* on the basis of utility yields the same ordering of households. Equivalence scales are typically defined as ratios of cost functions between demographically distinct households:

$$I\left(\mathrm{p}, \mathbf{z}, \mathbf{z}_0, u\right) = \frac{\mathrm{c}\left(\mathrm{p}, \mathbf{z}, u\right)}{\mathrm{c}\left(\mathrm{p}, \mathbf{z}_0, u\right)} \tag{2.1}$$

An equivalence scale measures the monetary cost of reaching utility u for a household of demographic composition $\mathbf{z}$ in proportion to the cost of reaching the same u for a baseline household with demographic composition $\mathbf{z}_0$, both facing the same vector of prices.[1]

We refer to u as a measure of welfare (or utility) throughout this chapter. The utility concept in applied welfare analysis requires a degree of *interpersonal comparability* of utility that goes further than the types of assumptions that are usually made in demand analysis. Interpersonal welfare comparability means that every household's welfare can be derived from one single (but not necessarily fully specified) utility

---

[1] Consider two households a and b. Both households face the same vector of prices p, have different household composition $\mathbf{z}^a$ and $\mathbf{z}^b$, and have nominal income levels $\mathbf{x}^a$ and $\mathbf{x}^b$ respectively. Household income $\mathbf{x}^a$ can be written as a (cost) function of the prices it faces, household composition $\mathbf{z}^a$ and the utility level it attains u^a:

$$\mathbf{x}^a = c\left(\mathrm{p}, \mathbf{z}^a, u^a\right) = c\left(\mathrm{p}, \mathbf{z}^0, u^a\right) \times \frac{c\left(\mathrm{p}, \mathbf{z}^a, u^a\right)}{c\left(\mathrm{p}, \mathbf{z}^0, u^a\right)} \tag{2.2}$$

where $\mathbf{z}^0$ is the household composition of a baseline household. Using equation (2.1), equation (2.2) can be rewritten:

$$\frac{\mathbf{x}^a}{I\left(\mathrm{p}, \mathbf{z}^a, \mathbf{z}^0, u^a\right)} = c\left(\mathrm{p}, \mathbf{z}^0, u^a\right) \tag{2.3}$$

One could do the same thing for household b:

$$\frac{\mathbf{x}^b}{I\left(\mathrm{p}, \mathbf{z}^b, \mathbf{z}^0, u^b\right)} = c\left(\mathrm{p}, \mathbf{z}^0, u^b\right) \tag{2.4}$$

Because $c\left(\mathrm{p}, \mathbf{z}^0, u^a\right)$ and $c\left(\mathrm{p}, \mathbf{z}^0, u^b\right)$ are monotonically increasing functions of u^a and u^b respectively, they can be used as a valid measure of utility itself. As a result, welfare comparisons between household a and b can be performed.

Further generalizations of this are possible. If for example, households are not only different in terms of demographics, but also face a *different* set of prices, nominal household income should be scaled with the following "adjusted" equivalence scale:

$$I\left(\mathrm{p}, \mathrm{p}^0, \mathbf{z}, \mathbf{z}^0, u\right) = \frac{\mathrm{c}\left(\mathrm{p}, \mathbf{z}, u\right)}{\mathrm{c}\left(\mathrm{p}^0, \mathbf{z}^0, u\right)} \tag{2.5}$$

function.[2] Our analysis depends on such a degree of interpersonal comparability of utility.[3] We do not however assume that u, our measure of utility, is cardinal. Any positive transformation of u (that does not depend on prices or demographics) would be an equally valid measure of utility.[4]

In this research we are estimating equivalence scales using a sample of Indonesian households. We put considerable emphasis on testing the price and reference utility dependency of equivalence scales. Equivalence scales that are typically used in practice are assumed to be functions of demographics only. In section (2.3) we decisively reject the null of independence of price as well as independence of utility (i.e., the equivalence scale's independency of utility is better known as independence of base (henceforth, IB) (Lewbel 1989)). We are able to reject independence of price and independence of base solely on the basis of analyzing patterns of demand over food and non-food items. In other words, the patterns of demand we observe in the data are inconsistent with household cost functions that produce equivalence scales that are independent of prices and independent of base. Therefore, we conclude that equivalence scales that are used in the Indonesian development practice, and that are merely functions of demographics, are inconsistent with observed patterns of demand. A potentially important issue is whether such equivalence scales also lead to inconsistent welfare comparisons.[5]

We have been subsequently interested in estimating the magnitude and direction of the price and utility dependence of equivalence scales. The empirical approach relies on the full specification and estimation of household cost functions. For identification we use data on self-rated consumption adequacy (a measure of self-rated welfare) alongside observed patterns of demands on food and non-food items [see e.g. Pollak and Wales (1979) and Blundell and Lewbel (1991) for a discussion on identification of equivalence scales]. We use data from the 2003 and 2004 waves of the consumption panel of the Indonesian Socio Economic Survey [henceforth SUSENAS, further data description follows in section (2.3)].

To our knowledge this is the first study that simultaneously analyzes the price *and* reference utility dependency of equivalence scales empirically. We show that recent empirical studies into the relationship between income, demographics and subjective

---

[2] Demand analysis typically relies on the supposition that preferences over (bundles of) goods, conditional on demographics are the same across households.

[3] Blundell and Lewbel (1991) argue that ..without interpersonal comparability of some sort estimating an equivalence scale is hopeless almost by definition, unless it were possible to observe the demands of a single household over goods as its demographic composition were changed."

[4] On could say that utility or welfare concept is "ordinal" in the sense that it cannot be used to produce statements like: 'household A experiences twice as much utility as household B experiences'. Instead, the welfare measure can be used to produce more realistic statements like: 'household A experiences more utility than household B'.

[5] Equivalent incomes may be inconsistent welfare measures when erroneous equivalence scales are used for its construction.

evaluations of welfare tend to oversimplify the utility function that is implicit from their empirical specifications. We discuss the limitations of three papers that employ subjective welfare data in order to estimate equivalence scales in section (2.3) (Kapteyn and Van Praag 1976, Pradhan and Ravallion 2000, Schwarze 2003). We show that these studies impose preference structures that imply homothetic demands for goods (i.e., Engel curves[6] are linear and pass through the origin) and weakly separable preferences between goods and demographics (i.e., Engel curves are independent of demographics). These (implicit) assumptions on preferences are not in accordance with the patterns of demand we see in 'real-life'. Moreover, under these assumptions, equivalence scales are independent of prices and independent of base by construction. We derive exact restrictions on preferences that produce these results and show how previous studies relate to these restrictive assumptions.

Homothetic preferences for goods and weakly separable preferences between goods and demographics are typically rejected on demand data (Deaton and Muellbauer 1980a, Ray 1983). For example, food budget shares tend to decrease in total expenditures and tend to increase in household size. Equivalence scales that are constructed on the basis of these assumptions therefore wrongfully assume at the outset that price and reference utility effects are absent. In this research we allow for non-homothetic preferences and non-separabilities in our empirical specification in order to test the empirical significance of both concepts.

We show that once we allow for sufficiently flexible preferences to accommodate more realistic patterns of demand, we need demand data alongside measures of self-rated welfare in order to obtain precise estimates of the preference parameters. We illustrate this efficiency issue within a Cobb-Douglas preference framework and by adopting a simple demographic parameterization of the preference parameters in section (2.3.3). We subsequently continue the analysis using more general functional form specifications (i.e., the almost ideal demand system) in section (2.3.4). We find that more elaborate parameterizations greatly outperform the simpler specifications in terms of statistical significance.

We find statistically significant price and reference utility effects. Equivalence scales increase in the food to non-food price ratio for low utility household, whereas we hardly see any price dependence for high utility households. An a priori unexpected finding is that equivalence scales tend to *increase* in reference utility, indicating that richer households demand a greater compensation (in relative terms) in response to a demographic expansion. Whereas these effects are highly important in statistical terms we find that they are arguably small in qualitative terms. Moreover, we conclude that the (subjective) data performs well by generating intuitive results. Each

---

[6] Engel curves of a certain good are defined as expenditures on that good, as a function of total expenditures.

additional child reduces the likelihood of a household's head perceiving household consumption as adequate for any given level of expenditure. The equivalence scales we estimate are of reasonable magnitude, and somewhat larger than the well-known modified OECD scales (De Vos and Zaidi 1997).

This chapter is organized as follows: in section (2.2) we discuss some of the literature and the limitations of the empirical literature on equivalence scales. Section (2.3) is the core of this research. We first introduce the data and report some of its important patterns in section (2.3.1). In section (2.3.2) we introduce a parametric model. In section (2.3.3) we discuss identification and estimation issues. Finally, in section (2.3.4) we estimate equivalence scales, and discuss their price and reference utility dependency.

2.2 The current literature and its limitations

Pollak and Wales (1979) show that equivalence scales are not identified from observed patterns of demands over goods. Observed household expenditures merely contain information on preferences for 'goods' q, conditional on household composition $\mathbf{z}$ (i.e., conditional preferences using their terminology), whereas equivalence scales depend on joint preferences over goods *and* household composition (i.e., unconditional preferences using their terminology). In other words, equivalence scales depend on the shape of indifference curves in $\mathrm{q}-\mathbf{z}$-space, whereas observed demands are merely informative about the shape of the indifference curves in q-space *conditional* on $\mathbf{z}$.

Blundell and Lewbel (1991) reiterate the claims made by Pollak and Wales (1979), but propose a less gloomy attitude towards the identification problem. They argue that if one is willing to agree on a scale in a base year (or price regime), demand data is sufficiently informative to identify equivalence scales in any other year (or price regime). Their result is mathematically represented by rewriting equation (2.1) in the following way:

$$\begin{aligned}\frac{c\left(\mathrm{p},\mathbf{z},u\right)}{c\left(\mathrm{p},\mathbf{z}^{0},u\right)} &= \frac{c\left(\mathrm{p},\mathbf{z},u\right)/c\left(\mathrm{p}^{0},\mathbf{z},u\right)}{c\left(\mathrm{p},\mathbf{z}^{0},u\right)/c\left(\mathrm{p}^{0},\mathbf{z}^{0},u\right)}\frac{c\left(\mathrm{p}^{0},\mathbf{z},u\right)}{c\left(\mathrm{p}^{0},\mathbf{z}^{0},u\right)} \\ &= \frac{L\left(\mathrm{p},\mathrm{p}^{0},\mathbf{z},u\right)}{L\left(\mathrm{p},\mathrm{p}^{0},\mathbf{z}^{0},u\right)}\frac{c\left(\mathrm{p}^{0},\mathbf{z},u\right)}{c\left(\mathrm{p}^{0},\mathbf{z}^{0},u\right)} \qquad (2.6)\end{aligned}$$

where $L\left(\mathrm{p},\mathrm{p}^{0},\mathbf{z},u\right)$ is a true-cost-of-living index that can be fully identified from demand analysis.[7] Equation (2.6) suggests at least three things. First, it suggest that

---

[7] The derivations of Blundell and Lewbel (1991) are insightful and clearly illustrate the identification issue: Demand equations $D\left(\mathrm{p},\mathbf{z},\mathbf{x}\right)$ derived from indirect utility function $S\left(\Psi\left(\mathrm{p},\mathbf{x},\mathbf{z}\right),\mathbf{z}\right)$ are the same as demand equations derived from $\Psi\left(\mathrm{p},\mathbf{x},\mathbf{z}\right)$.[8] Both preference functions yield the

equivalence scales depend on prices. Second, demand data contains useful information for estimating equivalence scales. The third suggestion is that demand analysis is sufficiently informative to test the null of price (in)dependency of equivalence scales. From studying demands therefore we may test the validity of some of the expert scales that are used in practice [e.g., the well-known (modified) OECD scales are independent of prices by construction (De Vos and Zaidi 1997)].

However, even if demands over goods are informative about equivalence scales, we still need to 'agree on a scale in a base year (or price regime) (Blundell and Lewbel 1991)' to estimate equivalence scales. That is, we need to solve the statistical identification problem put forward by Pollak and Wales (1979). In general, solving an identification problem requires arbitrary assumptions *or* additional data. The literature has done both. The well-known and widely used OECD scales exemplify the inherent arbitrariness often associated with equivalence scales. The OECD scales are so-called expert scales and are based on expert's intuition or experience. Experts form opinions about household specific needs and translate these into equivalence scales. The OECD experts conclude that if the costs for the first household member is normalized to 1, the cost of the second adult is 0.5 and any subsequent child costs 0.3 [see the "OECD modified" equivalence scales (De Vos and Zaidi 1997)].[9] From these cost assessments the OECD experts effectively pinpoint the shape of indifference curves in $\mathrm{q}-\mathbf{z}$-space.

It is important to note that other ways of determining equivalence scales rely on similar subjective cost assessments.[10] Olken (2005) for example, identifies equivalence scales on the basis of aid allocation at the local community level. Olken (2005) therefore assumes that local community authorities have sufficient insight in the needs of its people to make valuable judgements of welfare under different demographic regimes. Local community leaders, therefore, effectively determine the shape of the indifference curves in $\mathrm{q}-\mathbf{z}$-space.

In this research we rely on individual household heads themselves to rate their household specific needs [see for similar approaches e.g., Kapteyn and Van Praag (1976), Pradhan and Ravallion (2000) and Schwarze (2003)]. By means of direct questions on welfare levels across demographically distinct households we can recover

---

same ordering over goods q, such that, as a result, observed demands do not discriminate between the two preference functions. Consequently, the appropriate cost function, which is the inverse of indirect utility, cannot be retrieved from demands either. Household cost functions are said to be only partly identified by observed demand patterns (i.e., underidentification).

[9] The "OECD modified" equivalence scales associated with a two adults, two child family for example (where a one person household is the baseline) becomes $\frac{1+0.5+0.3+0.3}{1} = 2.1$.

[10] One could argue however that studying the simultaneous demand for goods and demographics identifies indifference curves in $\mathrm{q}-\mathbf{z}$ space, and thus equivalence scales. This is however practically impossible as many demographic variables are only partially subject to choice (e.g., children), others are not subject to choice (e.g., age).

certain features of indifference curves in $\mathrm{q}-\mathbf{z}$-space we need to construct equivalence scales.

Expert scales and scales based on individual assessments of welfare are not fundamentally different. The respective approaches differ in *who* is determining the cost of welfare under different demographic regimes, were it experts, local community leaders or household heads themselves. Surveying individuals about their welfare experiences is a way of gathering information on individual preferences that is increasingly gaining ground in economic research [see for example Kahneman and Krueger (2006)].[11] So far however, the literature on subjective measurement tends to rely on rather elementary empirical specifications. It is standard to assume a linear relationship between a measure of satisfaction/happiness and some measure of income and demographics (e.g., Blanchflower and Oswald (2004) out of many). This sharply contrasts to the empirical literature on demand which is much more advanced in terms functional relationships between variables (e.g., Banks et al. (1997)).

We show that overly simplistic functional form assumptions is one of the key limitations of the existing *empirical* literature on equivalence scales. Popular functional form assumptions produce equivalence scales that are independent of prices and IB by construction. Moreover, the behavioral responses that are implicit from these empirical specifications would not be accepted by the existing empirical literature on demand. Estimating the price and reference utility dependence of equivalence scales therefore, explicitly calls for reconciling the advanced literature on demand with the more *ad hoc* modeling style that is common in the literature on 'satisfaction'.

We will subsequently discuss some of the properties and limitations of three important papers in the field of equivalence scales [i.e., Schwarze (2003), Pradhan and Ravallion (2000) and the individual welfare function of income (henceforth, WFI) literature pioneered by Van Praag (1968)]. (The list could be easily extended with additional studies in the field of 'satisfaction' that exhibits similar linearity assumptions. Implications for demand functions are the same, and would be an interesting field for further research.) The definition of the variables in the subsequent examples below, are taken from the original studies.

1. Within the WFI literature households are asked to report which income levels match certain verbal qualifications (i.e., very bad, bad, insufficient, sufficient, good, excellent). These answers are supposed to represent some measure of utility where *very bad* gets assigned 1/12, *bad* gets assigned 3/12 etc. For example, Kapteyn and Van Praag (1976) subsequently fit a lognormal cumulative density function to relate the different income levels to these verbal qualifications.

---

[11] Individuals are typically interviewed to rate their own welfare experience, by choosing some verbal description of welfare that best describes their personal situation.

Their measure of utility or welfare u is defined as follows:

$$u = \Lambda\left(y;\mu,\sigma\right) = N\left(\ln y - \mu; 0, \sigma\right)$$

where

$$\mu = \beta_1 \ln fs + \beta_2 \ln y + \beta_3 + \varepsilon$$

$\ln y$ is measure of real income. $\ln fs$ is the natural logarithm of family size.

2. Schwarze (2003)'s model:

$$u = S^*_{it} = \beta_0 + \beta_1 \ln Y_{it} - \beta_1 e \ln h_{it} + X_{it}\beta_2 + PRICE\beta_3 + \eta_t\beta_4 + \varepsilon_{it}$$

where S^*_{it} is an (unobserved) ordinal measure of income satisfaction (that may be interpreted as utility u). h_{it} is household size (note the separability w.r.t. income). He uses a price index and the inflation rate for the variable $PRICE$.

3. Pradhan and Ravallion (2000): They postulate the probability of consumption adequacy as the probability of having expenditures above a subjective minimum z (this probability may be taken as a measure of utility u). The probability is a function of demographics ($\ln x$) and real expenditures ($\ln y$).

$$u = Pr\left(y > z\right) = F\left[\ln y - \alpha - \beta_1 \ln y - \beta_2 \ln x\right]$$

The three respective studies introduce indirect utility functions that can be written as cost functions simple algebraic manipulations. The cost functions may be subsequently written as follows:

$$c\left(\mathrm{p}, \mathbf{z}, u\right) = m_1\left(\mathbf{z}\right) b_1\left(\mathrm{p}\right) u \tag{2.7}$$

where $m_1\left(\mathbf{z}\right)$ is a function of demographic variables $\mathbf{z}$, $b_1\left(\mathrm{p}\right)$ is a function of prices, that is (or should be) homogenous of degree 1 in prices in order to satisfy adding up, and u is a measure of utility.

It is straightforward to infer from combining equation (2.1) and equation (2.7) that prices nor reference utility plays a role for equivalence scales in all three studies. Op top of that, one can show that under (2.7) preferences are homothetic and weakly separable between goods and demographics (i.e., food shares are merely functions of prices). Both assumptions are typically rejected in demand analysis as food shares tend to decrease in income and increase in household size.[12]

Note however that equivalence scales are independent of prices and reference

---

[12] To see this, apply Shephard's Lemma to the (log) cost function (equation (2.7)) to derive

utility for a much broader class of cost functions:

$$c\,(\mathrm{p}, \mathbf{z}, u) = m_2\,(\mathbf{z})\, b_2\,(\mathrm{p}, u) \tag{2.9}$$

where $m_2\,(\mathbf{z})$ a function of demographics, $b_2\,(\mathrm{p}, u)$ is a function of prices and utility and, again, is homogenous of degree 1 in prices.

Moreover, both equation (2.7) and (2.9) are special cases of an even broader class of cost functions known as the independence of base class (Lewbel 1989):

$$c\,(\mathrm{p}, \mathbf{z}, u) = m_3\,(\mathbf{z}, \mathrm{p})\, b_3\,(\mathrm{p}, u) \tag{2.10}$$

The IB properties, that is the parameter restriction that yield cost functions of the form described in (2.10), are tested against a more general alternative in the empirical section (2.3).

We know of three empirical studies that have used subjective data that allow for more general preference structure than equation (2.7). Van Praag (1991) and Kapteyn (1994) have extended the original WFI framework by allowing the parameters μ and σ to depend on prices in a way that it relates to the Almost Ideal Demand System (Deaton and Muellbauer 1980b). Van Praag (1991) however, does not estimate the parameters of the model in his paper. Perhaps a drawback of these studies is that the same parameters may be identified from both demand and subjective data. Typically, using subjective data one can identify more. Their assumption is relaxed in this research. Van Praag and Van der Sar (1988) also allow for more general preference structures, and are able to estimate base utility effects. Finally we would like to point to a study by Koulovatianos et al. (2005) that (implicitly) abandons assumption (2.9) and estimates the income dependence in an interesting non-parametric way.

2.3 The empirical analysis: towards estimating equivalence scales

In this section we work towards estimating the parameters of household cost functions that are subsequently used to construct equivalence scales. This calls for a

---

compensated food shares w_i:

$$w_i = \frac{\partial \ln c\,(\mathrm{p}, \mathbf{z}, u)}{\partial \ln \mathrm{p}_i} = \frac{\partial\,(\ln m_1\,(\mathbf{z}) + \ln b_1\,(\mathrm{p}) + \ln u)}{\partial \ln \mathrm{p}_i} = \frac{\partial \ln b_1\,(\mathrm{p})}{\partial \ln \mathrm{p}_i} \tag{2.8}$$

Because, compensated food shares do not depend on utility, compansated and uncompensated food shares are equal. Hence, Engel curves are linear, pass through the origin and independent of demographics.

few intermediate steps concerning data description/requirements and identification issues. First, in section (2.3.1) we elaborate on the data set and discuss some of its most interesting properties. We find that the data behaves remarkably well by exhibiting many intuitive features. Next, we introduce a parametric model in section (2.3.2) that can accommodate some of the important features of the data. Moreover, section (2.3.2) discusses the parameter restrictions that produce independence of price and independence of base within our parametric framework. Section (2.3.3) discusses the identification of the model parameters and shows the importance of utilizing demand data *and* subjective data in one integrated statistical framework for reasons of efficiency. Section (2.3.4) estimates equivalence scales using the almost ideal demand system preferences to govern demands. We subsequently present the estimated equivalence scales and discuss its price and reference utility dependency in section (2.3.5).

2.3.1 Data description and stylized facts

For this research we use a new and unexploited micro consumption panel data set drawn from Indonesian National Socioeconomic Survey (henceforth SUSENAS). The original SUSENAS is a representative survey among Indonesian households and interviews about 200,000 households on a yearly basis. Listings of consumption expenditure items of this broad survey however are rather limited. From 2002 up to 2004 the SUSENAS has selected a subset (about 10,000 households) of the original SUSENAS to take part in a consumption panel. In addition to a very detailed set of consumption expenditures and demographic variables, households were asked to rate their respective food and non-food consumption as adequate or inadequate for their particular household's needs (only the 2003 and 2004 wave). This data is referred to as self-rated consumption adequacy data.

Self-rated consumption adequacy has already been part in various surveys across developing countries. Pradhan and Ravallion (2000) for example use consumption adequacy data from Jamaica and Nepal to estimate subjective poverty lines. Questions on self-rated consumption adequacy are relatively easy to answer and could therefore be included in addition to the standard content of expenditure surveys without much effort.

The exact questions and answers to collect data on self rated consumption adequacy were phrased as follows:

Q1. In the past month, has your food consumption been adequate for your household needs?
1. no

2. yes
3. more than adequate
4. do not know

Q2. In the past month, has your non-food consumption been adequate for your household needs?
1. no
2. yes
3. more than adequate
4. do not know

We have grouped together the households that reported *yes* or *more than adequate* as the number of households reporting *more than adequate* is relatively small. Out of the 19.810 (2003 and 2004) households about 68% answered affirmatively (*yes* or *more than adequate*) to *Q1* are therefore food consumption adequate and more than adequate. About 27% reported *no* to *Q1* and around 4% did not know (or was otherwise missing). About 58%, 35% and 6% reported *yes/more than adequate*, *no* and missing to $Q2$ respectively. The households that have reported missing to either $Q1$ or $Q2$ have been eliminated from the analysis. Of the remaining households (i.e., 18,243 households) 60% reported *yes* or *more than adequate* to both questions and about 25% reported a *no* to both questions.

About 12% reported a *yes/more than adequate* to *Q1* and a *no* to *Q2*, where about 3% reported a *no* to *Q1* and a *yes/more than adequate* to *Q2*. Households that answer in this way (i.e., a *yes/more than adequate* and a *no*) comprise therefore a minority of the sample. However, we did not find any notable differences in observed demands where we expected them. One would for example expect that the food share in the total budget of 'food adequate, non-food inadequate' households is smaller than for other households. The group that is food adequate but non-food *in*adequate is sizable enough however, to be of qualitative importance such that there is room for studying these groups in more detail. This is however beyond the scope of this research and is therefore left for future research. For the analysis we use the observations that report two positive or two negative answers to *Q1* and *Q2* (=15525 observations). Henceforth we will talk about (total) consumption adequacy where consumption adequacy means both food and non-food adequate. Consumption *in*adequacy means both food and non-food inadequate.

When answering *Q1* and *Q2* it is likely that not all respondents had the same concept in mind. Still, we expect that there is some general consensus about what consumption adequacy means, and, that it should be related to total consumption in some way. Indeed, in the data we find that consumption adequate households consume about 50% more (in real terms) than households reporting inadequacy. (Real consumption constructed as total consumption expenditures, scaled with food poverty lines supplied by the World Bank. These food poverty lines measure the monetary cost of buying a minimum basket of food items in a particular region at a particular point in time.) This indicates that –at least on average– *Q1* and *Q2* measure what they are supposed to measure. If we further differentiate households on the basis of household size, we still find that consumption adequate household consume about 50% more than inadequate households.[13] These regularities strongly suggest that there is some general understanding –on average– about the concept of consumption adequacy.

In addition we expect that for a given level of total expenditures small families more often report consumption adequacy than large families. The three-way relationship between consumption, household size and consumption adequacy is nonparametrically displayed in table (2.1). Table (2.1) lists the frequency of consumption adequate households within different subgroups.[14] The subgroups are defined on the basis of real total expenditures and household size. Expenditure classes are presented on the vertical axis, and are defined by sorting total *real* household consumption[15] in five equally distant sections (i.e., group 1 consumes less than 4.4 times the poverty line for food, group 2 consumes between 4.4 and 8.8 times the poverty lines for food, etc.). Classes of household size are presented on the horizontal axis (size category 1 means a 1 or 2 person household. Size category 2 means a 3 or 4 person household. Size category 3 means a 5 or 6 person household and category 4 means 7 household members or more). We are consistently finding that conditional on real expenditures, family size decreases the frequency of consumption adequate households within our sample. Moreover, conditional on size, increasing total expenditures consistently increases the likelihood of reporting consumption adequacy.

One could read the frequencies reported in table (2.1) as points on an indirect utility function. A high frequency of adequate consumers within a certain group means high utility on average as a function of total expenditures and demographics. We should keep in mind however, that the expenditure categories are constructed independently of how total expenditures are built up. Because budget shares of food typically increase in household size, it is likely that the content of total consumption differs across categories of size. This is important for our purposes as we are interested in how price changes impact on equivalence scales. If for a given level of income, large households consume more food than small household, they will be differently affected by a change in food prices.

---

[13] We have looked at the same classes of household size as we adopt in table (2.1)

[14] That is the number of consumption adequate households divided by the total number of households within a group.

[15] total consumption divided by the poverty line for food

Table 2.1: *Frequency of consumption adequacy*

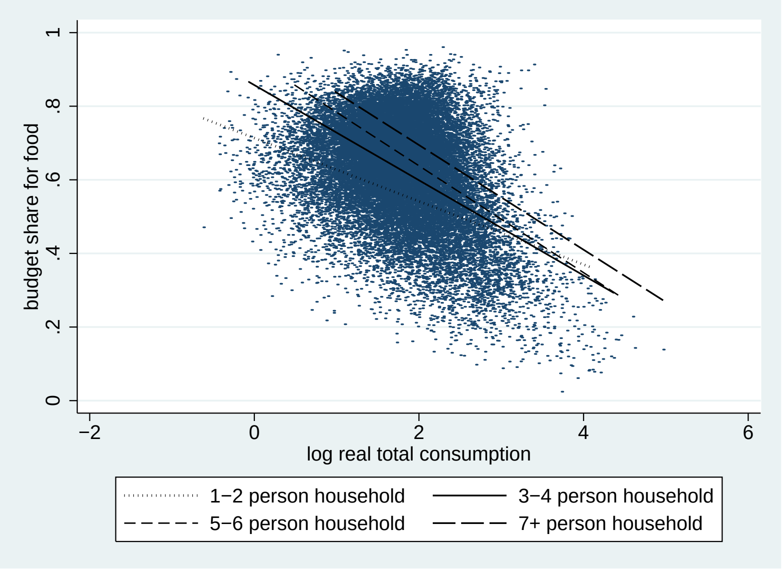
	Household size category				
Expenditures category	1	2	3	4	total
1	**0.61**	**0.50**	**0.48**	**0.26**	**0.54**
	1839	2014	337	34	4334
2	**0.86**	**0.71**	**0.66**	**0.53**	**0.69**
	821	3784	2075	546	7226
3	**0.97**	**0.90**	**0.82**	**0.73**	**0.84**
	111	987	955	413	2466
4	**1**	**0.95**	**0.92**	**0.85**	**0.92**
	35	267	378	155	835
5	**1**	**1**	**0.96**	**0.92**	**0.96**
	25	208	294	137	664
total	**0.70**	**0.69**	**0.72**	**0.67**	**0.7**
	2831	7260	4149	1285	15525

Note. The table reports the frequencies of households reporting consumption adequacy. Groups are discriminated on the basis of consumption levels (vertically) and household size (horizontally).

Figure (2.1) shows that the food consumption to total consumption ratio is decreasing in the *log* of real total expenditures, rejecting homothetic preferences (i.e., constant budget shares) for food and non-food. Sharp increases in food prices therefore, can be devastating for poor households as they spend roughly two thirds of their income on food. One of the key claims of our study is that previous studies explaining subjective measures of welfare tend to rely on rather elementary functional form specifications that are typically inconsistent with such expenditure patterns. Consequently, the simple preference structures that have been (implicitly) adopted predominantly in the satisfaction literature should be re-evaluated [see section (2.2)]. We have plotted four linear regression lines to discriminate the same classes of household size as we have used for constructing table (2.1). The figure shows that for a given level of expenditure, larger households spend more on food (a rival good) than on non-food (not rival goods per se). The pattern is more or less consistently revealed across categories of size. Yet, the regression lines seem to move closer together for higher levels of total consumption.

On the basis of this graph one can already conclude that household size matters for the allocation of total expenditures across food and non-food items, rejecting weakly separable preferences between consumption and demographics. Cost functions of the form (2.7) therefore are too restrictive to accommodate the features presented in figure (2.1). Section (2.3.2) introduces a parametric model that can accomodate these features of the data. Note that –at least in theory– the stylized facts of this section not necessarily imply that equivalence scales are dependent on prices or on reference utility. Under equation (2.9) for example, demand functions depend on demographics and total expenditures in some particular way.

Figure 2.1: *Budget share Engel curves for food*



Finally, we use food poverty lines as a region and time specific prices for food. The associated non-food prices are estimated using a method similar to Ravallion and Bidani (1994) [see appendix A.2].

2.3.2 A parametric model

Utility is an ambiguous concept as the term is used to describe both conditional and unconditional preferences [using the terminology of Pollak and Wales (1979)]. In this research we define the probability of being consumption adequate conditional on expenditures, prices and demographics as the concept of utility. Our equivalence scales therefore are "consumption adequacy" equivalence scales based on curves in $q - \mathbf{z}$-space that connect combinations of q and $\mathbf{z}$ that exhibit equal probability of consumption adequacy ($A = 1$ when consumption is adequate for household needs and $A = 0$ if consumption is not adequate for household needs).

$$u = S(\mathbf{x}, \mathbf{z}, \mathrm{p}, \eta, \theta) = P\left(A = 1 \middle| \mathbf{x}, \mathbf{z}, \mathrm{p}, \eta, \theta\right) \tag{2.11}$$

Note that equivalence scales are invariant to monotonically positive transformations of the *satisfaction* function S, as it would not affect the shape of the iso-utility curves in $\mathrm{q} - \mathbf{z}$-space. Analogous to the idea of utility maximization we assume that people are maximizing their consumption satisfaction by optimally allocating total expenditures over food and non-food items, for given prices and demographics.

η is a random effect that captures household specific unobserved characteristics affecting consumption adequacy. The parameters of the satisfaction function may be split up in two: $\theta = \{\theta^d, \theta^s\}$. θ^d is a vector of parameters that affects both demand and income satisfaction (but only through its impact on demand). θ^s is a vector of parameters that only affects satisfaction directly. Consequently, θ^d can be identified both from analyzing demand data and/or subjective data[16], whereas θ^s can be only identified from analyzing subjective data.

Suppose that demand equations can be derived from the following indirect utility function:

$$\Psi\left(\mathbf{x}, \mathbf{z}, \mathrm{p}, \theta^d\right) \tag{2.12}$$

By ordinality however, any positive monotone transformation of (2.12) yields the same demand equations. Utility –here modeled by the conditional probability of consumption adequacy– is therefore further specified as a positive monotone transformation of (2.12):

$$S\left(\Psi\left(\mathbf{x}, \mathbf{z}, \mathrm{p}, \theta^d\right), \mathbf{x}, \mathbf{z}, \mathrm{p}, \theta^s, \eta\right) = P\left(A = 1 \middle| \Psi\left(\mathbf{x}, \mathbf{z}, \mathrm{p}, \theta^d\right), \mathbf{x}, \mathbf{z}, \mathrm{p}, \theta^s, \eta\right) \tag{2.13}$$

Because both S and Ψ should yield the same demand equations we can show that $\Psi\left(\mathbf{x}, \mathbf{z}, \mathrm{p}, \theta^d\right)$ is a sufficient statistic for $\{\mathbf{x}, \mathrm{p}\}$ in S. This fact simplifies the analysis. Roy's identity, applied to both S and Ψ respectively should yield the same demand equation for any arbitrarily chosen good j:

$$\text{Roy's identity applied to } S: \qquad \mathrm{q}_j = -\frac{S_\Psi \Psi_{\mathrm{p}_j} + S_{\mathrm{p}_j}}{S_\Psi \Psi_{\mathbf{x}} + S_{\mathbf{x}}} \tag{2.14}$$

$$\text{Roy's identity applied to } \Psi: \qquad \mathrm{q}_j = -\frac{\Psi_{\mathrm{p}_j}}{\Psi_{\mathbf{x}}} \tag{2.15}$$

Both relations only equalize when the partial derivatives of S with respect to p and

---

[16] We show however that in practice, identifying these parameters from subjective data only, is difficult [see section (2.3.3)]

$\mathbf{x}$ are zero. Therefore, S depends only on p and $\mathbf{x}$ through its effects on Ψ. S, as a result, simplifies to:

$$S\left(\Psi\left(\mathbf{x},\mathbf{z},\mathrm{p},\theta^{d}\right),\mathbf{x},\mathbf{z},\mathrm{p},\theta^{s},\eta\right)=S\left(\Psi\left(\mathbf{x},\mathbf{z},\mathrm{p},\theta^{d}\right),\mathbf{z},\theta^{s},\eta\right) \tag{2.16}$$

Income, or price effects on consumption adequacy therefore pass only through Ψ. To establish the price dependency of equivalence scales therefore, it is of crucial importance to have precise estimates of the θ^{d} parameters *within* Ψ. Moreover, equation (2.17) shows that $\mathbf{z}$ affects utility (or consumption adequacy) through Ψ as well as directly.

For transparency of next section's derivations we further specify the utility function S. We should choose a specification for S that is able to accommodate the important features of demand we saw in section (2.3.1), i.e., non-linear Engel curves that depend on household size. The Almost Ideal Demand System is sufficiently flexible for this purpose and will be used in this study (Deaton and Muellbauer 1980a).[17]

We assume that the conditional probability to be consumption adequate relates to a normal CDF as follows:

$$\begin{aligned} u=S\left(\Psi\left(\mathbf{x},\mathbf{z},\mathrm{p},\theta^{d}\right),\mathbf{z},\theta^{s},\eta\right) &= \Phi\left(\Psi\left(\mathbf{x},\mathbf{z},\mathrm{p},\theta^{d}\right)+\delta\left(\mathbf{z}\right)+\eta\right) \qquad (2.17)\\ \Psi\left(\mathbf{x},\mathbf{z},\mathrm{p},\theta^{d}\right) &= \frac{\ln\mathbf{x}-\ln a\left(\mathrm{p},\mathbf{z}\right)}{b\left(\mathrm{p},\mathbf{z}\right)} \end{aligned}$$

where:

$$\begin{aligned} \ln a\left(\mathrm{p},\mathbf{z}\right) &= \alpha_0\left(\mathbf{z}\right)+\alpha_1\left(\mathbf{z}\right)\ln\mathrm{p}_1+\left(1-\alpha_1\left(\mathbf{z}\right)\right)\ln\mathrm{p}_2+0.5\gamma\left(\ln\frac{\mathrm{p}_1}{\mathrm{p}_2}\right)^2 \qquad (2.18)\\ b\left(\mathrm{p},\mathbf{z}\right) &= \exp\left[\beta_0\left(\mathbf{z}\right)\right]\left(\frac{\mathrm{p}_1}{\mathrm{p}_2}\right)^{\beta_1\left(\mathbf{z}\right)} \qquad (2.19) \end{aligned}$$

The α_0, α_1, β_0 and β_1 parameters are functions of a vector of demographic variables $\mathbf{z}$.[18] From this it is straightforward to derive a system of the two budget share equations for food and nonfood by an application of Roy's identity. Due to the adding up requirement one (of the two) budget shares is simply equal to one minus the other:

$$\mathbf{w} = \begin{pmatrix} w_1\left(\mathbf{x},\mathbf{z},\mathrm{p},\theta^{d}\right) \\ w_2\left(\mathbf{x},\mathbf{z},\mathrm{p},\theta^{d}\right) \end{pmatrix} = \begin{pmatrix} w_1\left(\mathbf{x},\mathbf{z},\mathrm{p},\theta^{d}\right) \\ 1-w_1\left(\mathbf{x},\mathbf{z},\mathrm{p},\theta^{d}\right) \end{pmatrix} \tag{2.20}$$

---

[17] Note, that we study only two composite goods, food and non-food consumption in this research.

[18] Further we define the function $\exp\left[\beta_0\left(\mathbf{z}\right)\right]$ instead of a more standard $\overline{\beta}_0\left(\mathbf{z}\right)$ as it matters for interacting the model parameters with $\mathbf{z}$ in a nontrivial way [see appendix (A.3)].

where $w_1\left(\mathbf{x},\mathbf{z},\mathrm{p},\theta^d\right)$ is the budget share for food.

$$\begin{aligned} w_1\left(\mathbf{x},\mathbf{z},\mathrm{p},\theta^d\right) &= \alpha_1\left(\mathbf{z}\right)+\gamma\ln\frac{p_1}{p_2} \\ &+\beta_1\left(\mathbf{z}\right)\left[\ln x-\alpha_0\left(\mathbf{z}\right)-\alpha_1\left(\mathbf{z}\right)\ln p_1-\left(1-\alpha_1\left(\mathbf{z}\right)\right)\ln p_2-0.5\gamma\left(\ln\frac{p_1}{p_2}\right)^2\right] \end{aligned} \tag{2.21}$$

The model is further operationalized by interacting the structural parameters with a vector of demographics $\mathbf{z}$. The way in which we do this is of fundamental importance for our purposes and basically lays down how equivalence scales are permitted to depend on prices and base utility. We assume that the preference parameters $\alpha_0\left(\mathbf{z}\right)$, $\alpha_1\left(\mathbf{z}\right)$, $\beta_0\left(\mathbf{z}\right)$, $\beta_1\left(\mathbf{z}\right)$ and $\delta\left(\mathbf{z}\right)$ are linear in the logarithm of household size $\ln\mathbf{z}$:

$$\begin{aligned} \alpha_0\left(\mathbf{z}\right) &= \alpha_0^0+\alpha_0^{\mathbf{z}}\ln\mathbf{z} \\ \alpha_1\left(\mathbf{z}\right) &= \alpha_1^0+\alpha_1^{\mathbf{z}}\ln\mathbf{z} \\ \beta_0\left(\mathbf{z}\right) &= \beta_0^0+\beta_0^{\mathbf{z}}\ln\mathbf{z} \\ \beta_1\left(\mathbf{z}\right) &= \beta_1^0+\beta_1^{\mathbf{z}}\ln\mathbf{z} \\ \delta\left(\mathbf{z}\right) &= \delta^0+\delta^{\mathbf{z}}\ln\mathbf{z} \end{aligned}$$

The price and reference utility dependence of equivalence scales

In this section we discuss the price and base utility dependence of equivalence scales within the context of our functional form specification. Given our functional form assumptions it is more convenient to write *log* cost functions or *log* equivalence scales, simply for notational purposes. The *log* cost functions are derived by inverting (2.17) with respect to the *log* of total expenditures $\ln\mathbf{x}$:

$$\ln\mathrm{c}\left(\mathrm{p},\mathbf{z},u\right)=\ln a\left(\mathrm{p},\mathbf{z}\right)+b\left(\mathrm{p},\mathbf{z}\right)\left[\Phi^{-1}\left(u\right)-\delta\left(\mathbf{z}\right)-\eta\right] \tag{2.22}$$

Equivalence scales are IB when the cost function (2.22) can be written as equation (2.10). Under our functional form assumptions equivalence scales are IB if $b\left(\mathrm{p},\mathbf{z}\right)$ is independent of demographic characteristics $\mathbf{z}$. Both β_0 and β_1 are linear functions of $\ln\mathbf{z}$ by assumption, such that $b\left(\mathrm{p},\mathbf{z}\right)$ becomes:

$$b\left(\mathrm{p},\mathbf{z}\right)=\exp\left[\beta_0^0+\beta_0^{\mathbf{z}}\ln\mathbf{z}+\beta_1^0\ln\frac{\mathrm{p}_1}{\mathrm{p}_2}+\beta_1^{\mathbf{z}}\ln\frac{\mathrm{p}_1}{\mathrm{p}_2}\ln\mathbf{z}\right] \tag{2.23}$$

$b\left(\mathrm{p},\mathbf{z}\right)$ is therefore independent of $\ln\mathbf{z}$ if and only if the following condition holds:

$$\beta_0^{\mathbf{z}}+\beta_1^{\mathbf{z}}\ln\frac{\mathrm{p}_1}{\mathrm{p}_2}=0 \tag{2.24}$$

This restriction on the parameters is tested in the empirical section (2.3).

In the empirical section we test the null of $\beta_0^{\mathbf{z}} = \beta_1^{\mathbf{z}} = 0$.[19] There have been many studies that reject the null of $\beta_1^{\mathbf{z}} = 0$, e.g., Kalwij et al. (1998), and we also find this in this research. β_1's dependency on demographics however, is in itself not conclusive to reject equation (2.24). A special case arises when IB holds at only at point on the $\frac{\mathrm{p}_1}{\mathrm{p}_2}$ domain. In that case IB is locally satisfied: equivalence scales do not depend on utility at one particular point on the price ratio domain. So, even though demand data suggests $\beta_1^{\mathbf{z}} \neq 0$, base dependency might not be much of an issue if $\beta_0^{\mathbf{z}} + \beta_1^{\mathbf{z}} \ln \frac{\mathrm{p}_1}{\mathrm{p}_2}$ is equal or close to zero in practice.

Perhaps a more interesting case to discuss is when IB is rejected (we reject IB with a with a test statistic of $\chi^2(2) = 76.86$, far greater than 13.82, the critical value associated with a 0.1% significance level). The obvious follow-up question is then whether equivalence scales increase or decrease in base utility *and* whether these effects are economically important. The magnitude of the base utility effects are not easily captured by a simple formula as it depends on all parameters, prices and utility itself. The economic importance of the base utility effects are discussed in the empirical section (2.3.5) where we report our final results. It is easier to derive expressions for the direction of the reference utility dependency of equivalence scales. $b(\mathrm{p}, \mathbf{z}) - b(\mathrm{p}, \mathbf{z}_0)$.[20] Given that $\mathbf{z} > \mathbf{z}_0$ it can be shown that equivalence scales depend positively on reference utility when:

$$\beta_0^{\mathbf{z}} + \beta_1^{\mathbf{z}} \ln \frac{\mathrm{p}_1}{\mathrm{p}_2} > 0 \tag{2.26}$$

And hence, given $\mathbf{z} > \mathbf{z}_0$ equivalence scales depend negatively on base utility when:

$$\beta_0^{\mathbf{z}} + \beta_1^{\mathbf{z}} \ln \frac{\mathrm{p}_1}{\mathrm{p}_2} < 0 \tag{2.27}$$

Within the context of the almost ideal demand system, the utility dependency of equivalence scales, therefore, relates to prices: we may have negative utility dependency in one price regime whereas we have positive utility dependency in another. In section (2.3.5) we go deeper into the interpretation of the base dependence condition $\beta_0^{\mathbf{z}} + \beta_1^{\mathbf{z}} \ln \frac{\mathrm{p}_1}{\mathrm{p}_2}$ when we discuss our results.

On the working of price effects on equivalence scales. Under standard regularity

---

[19] See appendix A.3 for an extensive exposition on price normalizations and demographic parameterizations and parameter testing in demand systems. This is important for our study, but for clarity we have moved this discussion to the appendix.

[20] The derivative of the logarithm of the equivalence scale with respect to utility u:

$$\frac{\partial}{\partial u} \ln I(\mathbf{z}, \mathbf{z}_0, \mathrm{p}, u) = \frac{\partial}{\partial u} \Phi^{-1}(u) \left[b(\mathrm{p}, \mathbf{z}) - b(\mathrm{p}, \mathbf{z}_0) \right] \tag{2.25}$$

conditions equivalence scales depend –if at all– on relative prices (i.e., price ratios). Cost functions are homogenous of degree 1 in prices such that equivalence scales –as a ratio of cost functions– are homogenous of degree 0 in prices. Hence, equivalence scales are not affected when all prices go up or down at the same rate. As a result of this property, equivalence scales formulas may be expressed in terms of price ratios. It can be subsequently derived that the derivative of the *log* equivalence scale with respect to the *log* price ratio is equal to the following:[21]

$$\frac{\partial \ln I\left(\mathbf{z}, \mathbf{z}_0, \frac{\mathrm{p}_1}{\mathrm{p}_2}, u\right)}{\partial \ln \frac{\mathrm{p}_1}{\mathrm{p}_2}} = w_1\left(\mathbf{z}, \mathrm{p}, u\right) - w_1\left(\mathbf{z}_0, \mathrm{p}, u\right) \tag{2.29}$$

where

$$w_1\left(\mathbf{z}, \mathrm{p}, u\right) = \alpha_1\left(\mathbf{z}\right) + \gamma \ln \frac{\mathrm{p}_1}{\mathrm{p}_2} + \exp\left[\beta_0\left(\mathbf{z}\right) + \beta_1\left(\mathbf{z}\right) \ln \frac{\mathrm{p}_1}{\mathrm{p}_2}\right]\left[\Phi^{-1}\left(u\right) - \delta\left(\mathbf{z}\right) - \eta\right]\beta_1\left(\mathbf{z}\right) \tag{2.30}$$

$w_1\left(\mathbf{z}, \mathrm{p}, u\right)$ is the compensated budget share equation for food (i.e., the first good). Hence, the price ratio elasticity of the equivalence scale is equal to the difference between the compensated food shares of demographically distinct households.

The equivalence scale's dependence on the price ratio is a function of many of the model parameters, utility and prices. It can be, however, inferred from equation (2.29) and (2.30) that equivalence scales are *globally* independent of prices when $\alpha_1^{\mathbf{z}} = \beta_1^0 = \beta_1^{\mathbf{z}} = 0$. Note that the independence of price condition for depends only on parameters that can be fully identified from demand analysis, which is a general property that follows from equation (2.6) (Blundell and Lewbel 1991).

2.3.3 Identification: Demand data and Subjective data in one econometric framework

The parameter vectors of interest θ^d and θ^s are estimated with maximum likelihood. Identification depends on a set of assumptions. First, we assume that the variables $\mathbf{x}$, $\mathbf{z}$ and p are exogeneous, in the sense that their (joint) marginal density function

---

[21]

$$\begin{aligned}
\frac{\partial \ln I\left(\mathbf{z}, \mathbf{z}_0, \mathrm{p}, u\right)}{\partial \ln \frac{\mathrm{p}_1}{\mathrm{p}_2}} &= \frac{\partial \ln I\left(\mathbf{z}, \mathbf{z}_0, \mathrm{p}, u\right)}{\partial \ln \mathrm{p}_1} \frac{\partial \ln \mathrm{p}_1}{\partial \ln \frac{\mathrm{p}_1}{\mathrm{p}_2}} = \frac{\partial \ln I\left(\mathbf{z}, \mathbf{z}_0, \mathrm{p}, u\right)}{\partial \ln \mathrm{p}_1} \frac{\partial \ln \mathrm{p}_1}{\partial \ln \mathrm{p}_1 - \ln \mathrm{p}_2} \\
&= \frac{\partial \ln I\left(\mathbf{z}, \mathbf{z}_0, \mathrm{p}, u\right)}{\partial \ln \mathrm{p}_1} \\
&= \frac{\ln \mathrm{c}\left(\mathrm{p}, \mathbf{z}, u\right) - \ln \mathrm{c}\left(\mathrm{p}, \mathbf{z}_0, u\right)}{\partial \ln \mathrm{p}_1} = w_1\left(\mathrm{p}, \mathbf{z}, u\right) - w_1\left(\mathrm{p}, \mathbf{z}_0, u\right) \qquad (2.28)
\end{aligned}$$

does not depend on θ^d and θ^s.

To derive the likelihood function we first define the a latent model for A and a stochastic version of the budget share equation (2.21). The latent model for A is effectively specified by equation (2.17) and can be written as follows:

$$A^* = \Psi\left(\mathbf{x}, \mathbf{z}, \mathrm{p}, \theta^d\right) + \delta\left(\mathbf{z}\right) + \eta + \varepsilon \tag{2.31}$$

Where η is a random effect, capturing individual specific unexplained persistence in the answers to the consumption adequacy question. ε is a time varying error term. The stochastic version of the budget share equation (2.21) is defined as follows:

$$w_1 = w_1\left(\mathbf{x}, \mathbf{z}, \mathrm{p}, \theta^d\right) + a + e \tag{2.32}$$

Where a is a random effect, that loosely allows for heterogeneity in preferences that is not accounted for by the model, and that stays constant over time. e is a time varying error term.

The stochastic properties of η, ε, a and e further specify the likelihood function. The first set of assumptions on the errors are:

$$\varepsilon, e \big| \eta, a, \varepsilon_{\bar{t}}, e_{\bar{t}}, \mathbf{x}, \mathbf{z}, \mathrm{p} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & \sigma_e^2 \end{bmatrix}\right) \tag{2.33}$$

where $\varepsilon_{\bar{t}}$ and $e_{\bar{t}}$ characterize future and lagged values of ε and e respectively. Without loss of generality, the variance of ε is normalized to one. If the random effects were known constants, we could now write down the individual specific likelihood function as follows:

$$\prod_t f\left(A, \mathbf{w}|\eta, a, \mathbf{x}, \mathbf{z}, \mathrm{p}, \theta\right) = \prod_t g\left(A|\eta, \mathbf{x}, \mathbf{z}, \mathrm{p}; \theta^d, \theta^s\right) \times h\left(\mathbf{w}|a, \mathbf{x}, \mathbf{z}, \mathrm{p}; \theta^d\right) \tag{2.34}$$

Because the random effects are inherently unobserved however, we need additional distributional assumptions on those. We assume that the random effects are indepedent of the exogenous regressors $\mathbf{x}$, $\mathbf{z}$ and p. It is however not unlikely that heterogeneity in one of the preference parameters is affecting consumption adequacy *and* the budget shares in the same way, such that we have allowed them to correlate:

$$\eta, a \big| \mathbf{x}, \mathbf{z}, \mathrm{p} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} \sigma_\eta^2 & \sigma_{\eta,a} \\ \sigma_{\eta,a} & \sigma_a^2 \end{bmatrix}\right) \tag{2.35}$$

The distributional assumptions on the random effects (2.35) are used to integrate them out of the likelihood (2.34). Consequently, the individual specific likelihood function that is used in estimation is the following:

$$L_i = \int_\eta \int_a \prod_t g\left(A|\eta, \mathbf{x}, \mathbf{z}, \mathrm{p}; \theta^d, \theta^s\right) \times h\left(\mathbf{w}|a, \mathbf{x}, \mathbf{z}, \mathrm{p}; \theta^d, \sigma_e^2\right) n\left(\eta, a; \sigma_\eta^2, \sigma_a^2, \rho_{\eta,a}\right) \mathrm{d}a\mathrm{d}\eta \tag{2.36}$$

where h is the conditional density associated with the budget share model, g is a conditional Bernoulli distribution associated with the binary outcome variable A, and n is the density of the random effects. The double integral is numerically approximated by means of Monte Carlo simulation. We derive the exact functional representation of the likelihood function (2.36) in appendix (A.1).

Because θ^d and θ^s are both present in g we could –at least in principle– rely on data on self reported food consumption adequacy, prices, demographics and total expenditures alone to estimate all necessary parameters. Using only the g part in the likelihood function, i.e., neglecting the element of the likelihood function containing h, sacrifices efficiency.[22] We show that for our purposes the efficiency issue is of paramount importance.

To highlight the added value of using demand data *alongside* subjective data to estimate parameters, we work out an example. We impose the following, admittedly overly restrictive, restriction on the parameters of equation (2.17): $\beta_1^0 = \beta_1^{\mathbf{z}} = \gamma_{11} = \delta^0 = \delta^{\mathbf{z}} = 0$. This parameterization produces Cobb-Douglas demands:

$$w_1 = \alpha_1(\mathbf{z}), \qquad w_2 = 1 - \alpha_1(\mathbf{z}) \tag{2.38}$$

Constant budget shares are obviously inconsistent with the data and we relax this assumption in section (2.3.4) [see figure (2.1)]. However, this parameter restriction relates to the empirical specifications that are typically adopted in the literature on satisfaction, such that it a relevant baseline case. Under the parameter restriction equation (2.17) collapses to the following:

$$u = \Phi\left(\exp\left[-\beta_0(\mathbf{z})\right]\left(\ln \mathbf{x} - \alpha_1(\mathbf{z}) \ln \mathrm{p}_1 - (1 - \alpha_1(\mathbf{z})) \ln \mathrm{p}_2 - \alpha_0(\mathbf{z}) + \eta\right)\right) \tag{2.39}$$

The parameters estimates of the Cobb-Douglas parameterization are reported in table (2.2), where the first three columns (1-3) are estimated using subjective data only (i.e., using the likelihood function specified by (2.37)). Column (4) estimates the same specification as column (3), but it utilizes demand data as an additional source

---

[22] Under the assumptions specified by (2.33) and (2.35) we could estimate the parameters of interest by relying only on subjective data and maximize the following (individual specific) likelihood function:

$$L_i^{\mathrm{SUBJECTIVE}} = \int_\eta \prod_t g\left(A|\eta, \mathbf{x}, \mathbf{z}, \mathrm{p}; \theta^d, \theta^s\right) n\left(\eta; \sigma_\eta^2\right) \mathrm{d}\eta \tag{2.37}$$

where $n\left(\eta; \sigma_\eta^2\right)$ is the marginal density of the random effects in the subjective model.

of variation (i.e., using the likelihood function specified by (2.36)).

Table 2.2: *Parameter estimates using Cobb Douglas preferences*

	(1) SUBJECTIVE	(2) SUBJECTIVE	(3) SUBJECTIVE	(4) SUBJECTIVE + DEMAND
$\alpha_0^{\mathbf{z}}$	0.585 (22.93)***	0.589 (22.83)***	0.564 (17.27)***	0.551 (13.16)***
α_0^0	0.488 (13.35)***	0.480 (12.91)***	0.509 (11.94)***	0.391 (7.40)***
$\alpha_1^{\mathbf{z}}$	–	-0.105 (-1.12)	-0.159 (-1.55)	0.029 (11.65)***
α_1^0	0.059 (1.10)	0.198 (1.48)	0.261 (1.88)*	0.572 (167.7)***
$\beta_0^{\mathbf{z}}$	–	–	0.061 (1.20)	0.116 (2.03)**
β_0^0	-0.426 (-12.80)***	-0.426 (12.80)***	-0.502 (-7.13)***	-0.393 (-5.14)***
$\ln\sqrt{\sigma_a^2 - \frac{\sigma_{\eta,a}^2}{\sigma_{eta}^2}}$	–	–	–	-2.277 (-151.3)***
$\ln\sigma_\eta$	0.206 (5.22)***	0.206 (5.22)***	0.207 (5.24)***	0.161 (3.99)***
σ_e	–	–	–	-2.403 (-249.2)***
$\tau = \frac{\sigma_{\eta,a}}{\sigma_\eta}$	–	–	–	0.041 (16.18)***
observations	15516	15516	15516	15516
log likelihood	-7944	-7943	-7943	1681

NOTE. The first three columns (1-3) estimates only employ the consumption adequacy data to estimate parameters. Column (4) uses both consumption adequacy data and the variation in the demands to estimate parameters. The nuisance parameters σ_η, σ_a, $\sigma_{\eta,a}$ and σ_e are transformed in estimation. The nonzero parameter τ implies that there is correlation between the two random effects η and a.
z statistics in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

The column (1) results report the estimates when only α_0 is interacted with *log* household size (i.e., $\beta_1^0 = \beta_1^{\mathbf{z}} = \gamma_{11} = \delta^0 = \delta^{\mathbf{z}} = \alpha_1^{\mathbf{z}} = \beta_0^{\mathbf{z}} = 0$). We find that the estimated $\alpha_0^{\mathbf{z}}$ is positive. This means that conditional on total expenditures larger households are less likely to report that consumption is adequate for household needs, a result that is consistent with other findings in the literature [see e.g., Pradhan and Ravallion (2000)].[23] This specification corresponds to the type of empirical models that are typically estimated on subjective data and imply equivalence scales that are independent of prices and IB [see most of the papers of the WFI literature, Pradhan and Ravallion (2000) and Schwarze (2003) for examples]. Here, (a positive transformation of) a measure of subjective welfare depends linearly and additively on a measure of real income and on (a vector) of demographic variables. In section (2.2) we have shown that this type of models exhibit homothetic preferences

---

[23] Note that the exponential transformation of the β_0 parameter restricts the effect of total consumption on consumption adequacy on consumption to be nonnegative. This is not a *binding* restriction given our data however, as the effect tends to be positive.

(budget shares are constant) and separable preferences between demographics and consumption (budget shares are independent of demographics), which are properties of demand that are highly unlikely given the patterns we see in figure (2.1).

In column (2) and (3) we also allow both $\alpha_1(\mathbf{z})$ and $\beta_0(\mathbf{z})$ to depend on *log* household size. The $\alpha_1(\mathbf{z})$'s dependence on household size would reject that equivalence scales are independent of prices and $\beta_0(\mathbf{z})$'s dependence on household size would reject IB. Instead, we find that the likelihood increases from both extensions are marginal suggesting that equivalence scales are independent of prices *and* IB.

We reach however, an entirely different conclusion if use demand data as an additional source of variation in estimation (i.e., using the likelihood function (2.36)). It is clear that observed patterns of demand are highly informative about $\alpha_1(\mathbf{z})$ the only parameter in Ψ in this specification. Contrasting the results of column (3), we find that $\alpha_1(\mathbf{z})$ depend positively and significantly on *log* household size in column (4). This captures the well-known empirical regularity that large households spend a larger fraction of total consumption on food. Consequently, this finding suggests that prices *do* matter for equivalence scales. At least within this particular parameterization: holding all else equal, an increase in food prices increases equivalence scales [see equation (2.29)].[24] In column (4) we also find that $\beta_0(\mathbf{z})$ significantly depends on household size which rejects IB [see equation (2.24)].

These preliminary results clearly indicate that even within relatively crude preference structures, excluding variation in demands is problematic for testing the price and utility dependence of equivalence scales. Where subjective data theoretically *identifies* all the important parameters of the cost function, it is insufficiently *powerful* to do this with reasonable precision. Hence, based on the column (1-3) results we would, spuriously, conclude that equivalence scales are independent *both* of prices and reference utility. Using demands as an additional source of variation we conclude the opposite.

The results of table 2.2 clearly motivate the empirical and theoretical relevance of our study because prices and utility seem to matter for equivalence scales. On top of that, we show that the standard empirical specification that is used in the satisfaction (or happiness) literature – the column (1) type model– is not in accordance with demand behavior.

2.3.4 An Almost Ideal Demand System

From figure (2.1) we conclude that the assumption of constant budget shares is too strict. As opposed to the restrictive parameterization we have discussed in the

---

[24] Intuitively: After a demographic expansion, households allocate a greater fraction of total expenditures on food. When food prices are high, compensating these households for a the loss in utility is more costly, such that equivalence scales increase in food prices.

previous section, the almost ideal demand system allows for more elaborate budget share equations and hence, also for more elaborate formulations of equivalence scales. We report estimates of three different regression models in table (2.3). The first column employs only subjective data in estimation and once again indicates the efficiency problems. The z-statistics are very low indicating that parameters are imprecisely estimated. The issue is resolved in the column (2) and (3) where we use demand data as an additional source of variation.

Table 2.3: *Parameter estimates using Almost Ideal preferences*

	(1) SUBJECTIVE	(2) SUBJECTIVE + DEMAND	(3) SUBJECTIVE + DEMAND
$\alpha_0^{\mathbf{z}}$	0.819 (0.42)	0.654 (7.34)***	0.654 (18.02)***
α_0^0	7.530 (0.17)	1.078 (0.94)	1.667 (8.47)***
$\alpha_1^{\mathbf{z}}$	1.675 (0.11)	0.017 (0.53)	–
α_1^0	-1.291 (-0.11)	0.588 (6.16)***	0.539 (28.75)***
$\beta_1^{\mathbf{z}}$	0.252 (0.37)	-0.030 (-8.77)***	-0.029 (-9.10)***
β_1^0	-0.207 (-0.42)	-0.081 (-17.66)***	-0.082 (-18.99)***
γ	2.473 (1.47)	-0.146 (-8.75)***	-0.137 (-31.37)***
$\beta_0^{\mathbf{z}}$	0.048 (0.42)	0.041 (0.85)	0.041 (0.91)
β_0^0	-0.519 (-3.68)***	-0.506 (-7.59)***	-0.506 (-7.92)***
$\delta^{\mathbf{z}}$	-0.136 (-0.03)	0.039 (0.22)	–
δ^0	11.887 (0.16)	0.998 (0.53)	1.975 (7.72)***
$\ln\zeta = \ln\sqrt{\sigma_a^2 - \frac{\sigma_{\eta,a}^2}{\sigma_\eta^2}}$	–	-2.766 (-128.1)***	-2.765 (-128.2)***
$\ln\sigma_\eta$	0.189 (4.74)***	0.149 (3.75)***	0.149 (3.76)***
$\ln\sigma_e$	–	-2.438 (-256.9)***	-2.438 (-257.0)***
$\tau = \frac{\sigma_{\eta,a}}{\sigma_\eta}$	–	0.016 (10.19)***	0.016 (10.18)***
observations	15516	15516	15516
log likelihood	-7904	4851	4851

NOTE. The first column estimates only employ the consumption adequacy data to estimate parameters. Column (2) and (3) use both the consumption adequacy data and the variation in the demands to estimate parameters. Column (3) restricts the insignificant parameters of column (2) to zero (except for β_0 *log* hhs because this parameters is not uniquely identified [see appendix A.3 for details]. The nuisance parameters σ_η, σ_a, $\sigma_{\eta,a}$ and σ_e are transformed in estimation. The nonzero parameter τ implies that there is correlation between the two random effects η and a.
z statistics in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

It follows immediately from the results column (2) of table (2.3) and column (4) of table (2.2) that the extension from Cobb Douglas to AIDS is a statistically significant improvement of fit. The log-likelihood increases from 1681 to 4851 while the number estimated parameters goes up by 5: $LR = 2 \times (4851 - 1681) = 6340$. The test statistic greatly exceeds the critical value of 20.52 that is associated with a 0.1% significance level, and rejects Cobb-Douglas in favor of AIDS. The estimated β_1 parameter is significantly negative, and, that β_1 is significantly *more* negative for larger households. In other words, food is a necessary good and food is even 'more necessary' for larger households, reflecting the rival nature of food.

In column (3) we restrict the insignificant parameters $\alpha_1^{\mathbf{z}}$ and $\delta^{\mathbf{z}}$ to zero.[25] The parameter restriction hardly affects the likelihood. The column (3) results are used for constructing equivalence scales in the subsequent section.

As a side-step we would like to mention that the simultaneous use of demand data and subjective data in an integrated econometric framework offers an opportunity for testing the overidentifying restrictions on the θ^d parameters. In this paper these restrictions are imposed to increase efficiency.[26] Kapteyn (1994) for example, argues the following: 'The subjective measures fully identify cost functions and the expenditure data do this partly. This makes it possible to test the null hypothesis that both types of data are consistent with one another, i.e., that they measure the same thing'. We have attempted to perform such a test for this research as well. However, due to the relative complex functional (AIDS) and distributional assumptions (notably the correlation between the two random effects) the likelihood routine had problems finding the optimum. Solving this practical obstacle is not straightforward and, we argue, is beyond the scope of this study. Nevertheless, studying this is a very interesting and important area of future research.

Finally, homogeneity and Slutsky symmetry are imposed. Negativity is not a priori imposed and has been tested. We evaluated the signs of the eigenvalues of the Slutsky matrix for every observation in sample conditional on the estimated parameters and we found no problems there (Deaton and Muellbauer 1980b).

---

[25] Note that $\beta_0^{\mathbf{z}}$ is insignificantly zero as well. Yet, we are not restricting this parameter to zero. Its size and significance level depend on arbitrary price normalizations [see appendix A.3]. Hence, the $\beta_0^{\mathbf{z}}$ parameter is not uniquely identified and may be negative or positive, significant or insignificant simply by choosing a particular price normalization vector $\{k_1, k_2\}$ [see Appendix A.3 for definitions of k_1 and k_2].

[26] These tests are interesting for at least two reasons. First, if both subjective and demand data *do not* measure the same thing the use of subjective data in studying economic behavior becomes ambiguous. Second, if *they do* measure the same thing, both sources of information are substitutes and can be used interchangeably (whatever source is available). Subjective data has the additional advantage that it does not suffer from the identification problem put forward by Pollak and Wales (1979).

Testing the price and utility dependence of equivalence scales

We are subsequently interested in testing the independence of base restriction. The IB condition (2.24) is satisfied when both $\beta_0^{\mathbf{z}}$ and $\beta_1^{\mathbf{z}}$ are zero. Note that condition (2.24) holds also in the 'unique' situation when $\beta_0^{\mathbf{z}}$ and $-\beta_1^{\mathbf{z}} \ln \frac{\mathrm{p}_1}{\mathrm{p}_1}$ are in balance. This possibility should not be ruled out a priori, but has (perhaps) only some local importance as price ratios differ widely across regions. We test the null hypotheses of IB:

$$H_0 \quad : \quad \beta_0^{\mathbf{z}} = \beta_1^{\mathbf{z}} = 0 \tag{2.40}$$

$$H_{\mathrm{a}} \quad : \quad \text{otherwise} \tag{2.41}$$

The associated test statistic is $\chi^2(2) = 76.86$ and rejects IB at a 0.1% significance level. This test shows that utility is (at least statistically) an important argument of equivalence scales.

Equivalence scales are globally independent of prices if $\beta_1(\mathbf{z}) = 0$ and $\alpha_1(\mathbf{z}) = \alpha_1$:

$$H_0 \quad : \quad \beta_1^0 = \beta_1^{\mathbf{z}} = \alpha_1^{\mathbf{z}} = 0 \tag{2.42}$$

$$H_{\mathrm{a}} \quad : \quad \text{otherwise} \tag{2.43}$$

Note, that to test the global independence of prices we need only to rely on demand analysis (Blundell and Lewbel 1991). We have argued before that the null hypothesis of global independence of price does not reflect reality if non-linear, household specific Engel curves are important properties of expenditure data. The test statistic associated with the null hypothesis above is $\chi^2(3) = 4383.17$ and rejects independence of price at a 0.1% significance level (the critical value at 0.1% significance is 16.27). This is clear evidence against the independence of price assumption and suggests important price dependence of equivalence scales.[27]

Under the null of global independence of prices, relative price changes do not impact on equivalence scales because household size does not affect compensated budget shares [see equation (2.29)]. However, even under the alternative hypothesis equivalence scales may be locally independent of prices. This depends on a range of factors, such as the price ratio itself, utility and demographics. We study the quantitative or economic importance of price and reference utility effects in the subsequent section (2.3.5).

---

[27] The parameters that we restrict under the above null are independent of arbitrary price normalizations [see appendix (A.3)].

2.3.5 Equivalence scales as functions of prices and reference utility

In the previous section we have provided statistical evidence for the importance of the utility and price ratio dependence of equivalence scales. In this section we discuss the economic relevance of these effects. Using the parameter estimates of column (3) of table (2.3) and the observed range of the price ratio and utility (the condition probability of reporting consumption adequacy) we can calculate equivalence scales as functions of prices and utility on meaningful domains. In figure (2.2) we show the empirical densities of the food/non-food price ratios that are used in estimation and the predicted empirical density of $\Phi^{-1}(u)$ conditional on the data set and the parameter estimates from column (3) of table (2.3). For calculating equivalence scales we impose that the individual random effects η are zero.

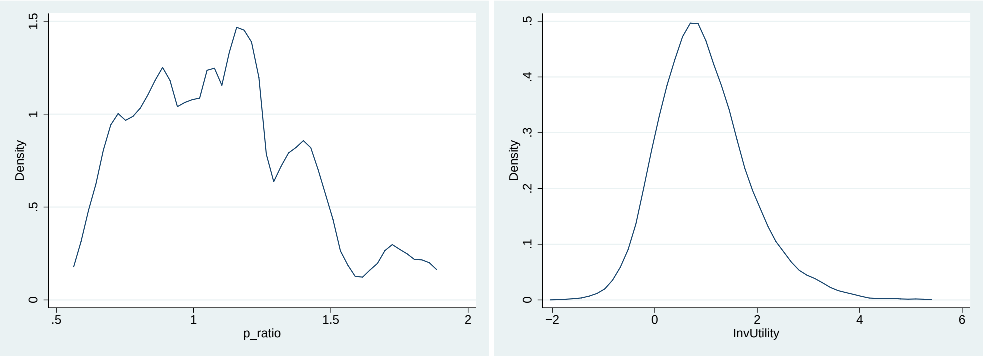


Figure 2.2: *Empirical densities. The left picture is the distribution of the food/non-food price ratios used in estimation. The right figure is empirical distribution of the predicted "$\Phi^{-1}(u)$" conditional on the data set and the parameter estimates from column (3) of table (2.3), where u is the predicted probability of reporting consumption adequacy.*

The four panels in figure (2.3) report equivalence scales (the curves) as functions of the food to non-food price ratio, and utility, for 2-, 3-, 4-, and 5- person households respectively. A one-person household is used as a baseline. The dashed-dotted curves are equivalence scales for high utility households (the 9th decile of the predicted utility distribution of figure (2.2)). The solid curves are equivalence scales for middle utility households (the 5th decile of the utility distribution). The dashed curves are equivalence scales for low utility households (the 1th decile of the utility distribution). The thick straight line in each panel is the modified OECD scale that is used as a benchmark (De Vos and Zaidi 1997).

The general impression of figure (2.3) is: *1.* The the scales are of reasonable mag-

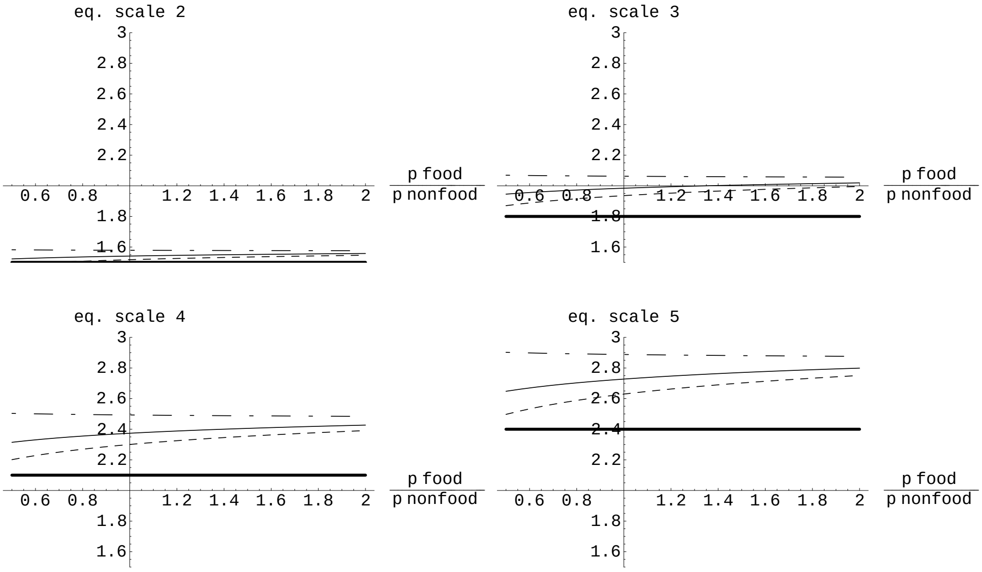


Figure 2.3: *The figure presents equivalence scales as functions of prices and reference utility for 2-, 3-, 4-, and 5- person households (a one person household is used as the baseline). The upper-left panel presents the equivalence scale for a 2 person household (upper-right panel (3 person), lower-left panel (4 person), lower-right (5 person)). For each plane, the food to non food price ratio is on the horizontal axis. Within each panel, the top, middle and bottom curves always represent equivalence scales for high, middle and low utility households (we use the 1st, 5th and 9th decile). The thicker straight line represents the modified OECD scale (De Vos and Zaidi 1997).*

nitude. They are in the same ballpark as the modified OECD scales. However, they are a little larger. This is potentially important as using the modified OECD scales may overstate the welfare of larger households when compared to smaller households. *2.* Equivalence scales *increase* in utility. *3.* Equivalence scales increase in the food/nonfood price ratio for low utility households, whereas the scales for high utility households are practically invariant to price changes. *4.* Whereas the price and utility dependence of equivalence scales are statistically significant, one may argue that in qualitative terms, the dependence is not very important.

The overall magnitudes of the estimated equivalence scales emphasize that subjective welfare data is useful for estimating equivalence scales (in the sense that they mirror the experts opinions rather well). We therefore conclude that both experts and individual Indonesian household heads –at least on average– have a similar idea about consumer needs across household types. Nevertheless, the OECD scales seem to underestimate the size of the scales for all household types, an effect that becomes more apparent in the bottom two panels of figure (2.3). Approximately, we estimate an equivalence scale of 1.5 for a two person household (OECD modified scale: 1.5), 2.0 for a three person household (OECD modified scale: 1.8), 2.4 for four person households (OECD modified scale: 2.1) and for a five person household 2.7 (OECD modified scale: 2.4).

In the remainder of this section we will discuss the economic forces driving two important features of household equivalence scales: *1.* equivalence scales depend positively on base utility on the price domain that is used in estimation. *2.* equivalence scales increase in the food/non-food price ratio for low utility households and are practically flat (or decrease marginally) on the food/non-food price ratio domain for high utility households.

It is illustrative to write down the *log* equivalence scales formula conditional on the parameter restriction $\alpha_1^{\mathbf{z}} = \delta^{\mathbf{z}} = 0$ (these parameters are insignificant in estimation and restricted at zero for constructing the scales):

$$\begin{aligned} \ln I\left(\mathrm{p}, \mathbf{z}, \mathbf{z}_0, u\right) = {} & \alpha_0\left(\mathbf{z}\right) - \alpha_0\left(\mathbf{z}_0\right) + \\ & \left(\exp\left[\beta_0\left(\mathbf{z}\right) + \beta_1\left(\mathbf{z}\right)\ln\frac{\mathrm{p}_1}{\mathrm{p}_2}\right] - \exp\left[\beta_0\left(\mathbf{z}_0\right) + \beta_1\left(\mathbf{z}_0\right)\ln\frac{\mathrm{p}_1}{\mathrm{p}_2}\right]\right) \times \\ & \left(\Phi^{-1}\left(u\right) - \theta^0 - \eta\right) \end{aligned} \tag{2.44}$$

One can interpret the formula of the *log* equivalence scale above as consisting of two separate elements. A *fixed* element[28] that is invariant to changes in prices and utility, and a *variable* element[29] that is a function of prices and utility. The interaction between prices and utility indicate once more that price effects and utility effects cannot be considered in isolation.

Discussing dependence of base. In figure (2.3) we find that equivalence scales depend positively on base utility. That is, we find that equivalence scales *increase* in utility for all observations (i.e., for all price ratios). The positive utility dependency of equivalence scales is a result of $\beta_0^{\mathbf{z}} + \beta_1^{\mathbf{z}} \ln \frac{\mathrm{p}_1}{\mathrm{p}_2} > 0$. This means that holding utility constant in response to a demographic expansion $\mathbf{z}_0 \to \mathbf{z}$ is more costly –in relative terms– for high utility households. The implications of this finding may be a bit perverse as it seem to suggests that rich households, for example, should receive more child allowance than poor households, both in relative as in absolute terms. The legitimacy of the argument however, depends on whether governments take up the objective to keep utility constant after children are born, irrespective of whether

---

[28] $\alpha_0\left(\mathbf{z}\right) - \alpha_0\left(\mathbf{z}_0\right)$

[29] $\left(\exp\left[\beta_0\left(\mathbf{z}\right) + \beta_1\left(\mathbf{z}\right)\ln\frac{\mathrm{p}_1}{\mathrm{p}_2}\right] - \exp\left[\beta_0\left(\mathbf{z}_0\right) + \beta_1\left(\mathbf{z}_0\right)\ln\frac{\mathrm{p}_1}{\mathrm{p}_2}\right]\right) \times \left(\Phi^{-1}\left(u\right) - \theta^0 - \eta\right)$

the household is rich or poor. Perhaps a more defensible government strategy would be to transfer money from the rich to the poor to promote equality. After all, why would rich households need financial compensation for an additional child, even if they experience large drops in welfare after having them? In fact, we would argue that they do not. There seems to exist however, a widespread political and scientific belief that equivalence scales should depend negatively on base utility.

Additionally, the utility dependency of equivalence scales is dependent on prices through our estimate for $\beta_1^{\mathbf{z}}$. $\beta_1^{\mathbf{z}}$ is smaller than zero and indicates that food consumption is more necessary for large households than for small households.[30] Large households as a result, substitute from food to non-food more rapidly as utility (or income) increases. It is this property of demand that causes the size and the direction of the base utility dependence to depend on the price ratio. An increase in the food/non-food price ratio mitigates the magnitude of $\beta_0^{\mathbf{z}} + \beta_1^{\mathbf{z}} \ln \frac{\mathrm{p}_1}{\mathrm{p}_2}$ and may even cause the condition to be negative such that –when food prices are high enough– equivalence scales would depend negatively on base utility. Our finding may be explained by the fact that the food/non-food price ratio is simply too low to yield a pattern of base dependency that is consistent with the general political and scientific consensus. It should be interesting to see whether we find similar results using data from developed countries.

Discussing price dependence. For price effects we find that equivalence scales increase substantially for low utility households whereas for high utility households equivalence scales are practically invariant to price changes. Hence, price effects and reference utility effects are highly interrelated within this framework. From equation (2.17) we know that price changes only affect equivalence scales through its impact on demands. By an application of Shephard's Lemma to the *log* equivalence scale formula we relate the price dependency of equivalence scales rather straightforwardly to compensated or Hicks demand equations [see equation (2.29)]

A marginal change in food prices therefore does not affect equivalence scales, if (and only if) the compensated food shares are not, or hardly affected by the demographic expansion $\mathbf{z}_0 \to \mathbf{z}$. Apparently, for high utility households this is exactly what happens. For low utility households, compensated food shares increase in response to a demographic expansion. The mechanism works as follows. Conditional on total expenditures $\mathbf{x}$, households will substitute towards food in an optimal response to a demographic expansion (i.e., a demographic effect). Because there are more mouths to feed, utility (as the probability of consumption adequacy) will decrease. Financial compensation for this drop in utility, in turn, motivates households to decrease food shares (i.e., an income effect). We find that for high utility households both the

---

[30] the combined $\beta_1(\mathbf{z})$ is smaller than zero $\forall \mathbf{z} > 0$, such that for all household types food is a necessary good.

demographic effect and the income effect are in balance. For low utility households this is different. For this type of households the demographic effect is stronger than the income effect. Equivalence scales for low utility households therefore increase in food prices.[31]

We present the working of the mechanism in figure (2.4). Figure (2.4) presents predicted uncompensated budget shares for food as functions of total expenditures. In both panels the top curve reflects the predicted budget shares for four person households and the bottom curve reflects the predicted budget shares for two person households. Point A links income levels of a two person, low utility household to its predicted food share. When this household grows to become a four person household it will consume more food conditional on total expenditures and moves up vertically to point B. When this household is subsequently receiving financial compensation for the loss of utility when moving from A to B it consequently slides along the curve to C. The vertical difference between C and A is the difference between the compensated food shares for large and small households and shows that an increase in food prices would increase equivalence scales for low utility households. A similar story applies to high utility households where a two person household goes from D to E to F sequentially. The vertical difference between F and D is practically zero. Price changes therefore affect equivalence scales only marginally for high utility households.

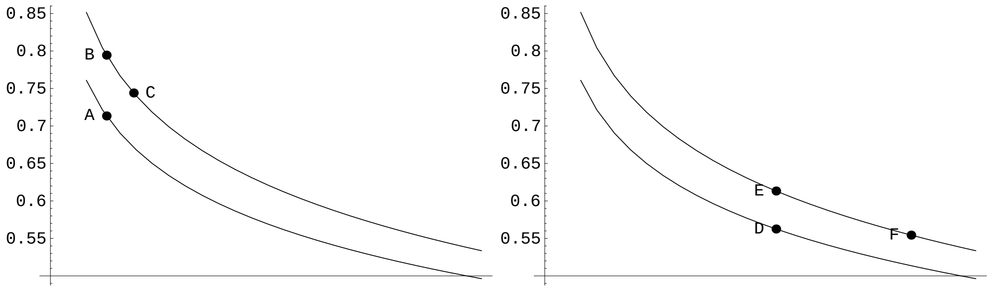


Figure 2.4: *The figure shows predicted* uncompensated *food shares as functions of total expenditures. The top curves reflect predicted food shares for large (four person) households. The bottom curves are the predicted food shares for small (two person) households. A, B and C are demands for low utility households. E, D and F are demands for high utility households.*

---

[31] Equivalence scales also increase in the food/non-food price ratio and decrease in non-food prices.

2.4 Summary and conclusion

The legitimacy of applied welfare analysis depends fundamentally on sound measures of welfare. To construct measures of welfare, the common strategy in the policy practice is to scale household income with a household specific index number. An index number that typically depends on household size. These index numbers are known as equivalence scales if and only if equivalent incomes (incomes that are constructed by scaling) are proper measures of utility, that is, ranking households on the basis of equivalent income yields the same ranking as ranking households on the basis of utility. In this research we specify and subsequently estimate these indices.

It has long been recognized that equivalence scales cannot be identified from observed patterns of demand (Pollak and Wales 1979). At least, if one agrees to rule out the possibility of estimating the demand for demographics and goods simultaneously. To construct equivalence scales one needs information about the shape of the indifference curves in $\mathrm{q} - \mathbf{z}$ (i.e., goods − demographics) space. Since demand data is merely informative about the shape of indifference curves in q space conditional on $\mathbf{z}$, we need additional sources of variation to solve this identification problem. We utilize self-rated consumption adequacy data drawn from the Indonesian SUSENAS consumption panel (2003-2004) for this purpose. Self-rated consumption adequacy data is informative about indifference curves in $\mathrm{q} - \mathbf{z}$ space and can be used to estimate equivalence scales. The SUSENAS furthermore, records detailed information on consumer expenditures and demographics.

Three seminal papers in this field discuss the potential impact of base utility and prices on these scales (Pollak and Wales (1979) by implicitly by discussing demand effects, Lewbel (1989) and Blundell and Lewbel (1991)). The empirical evidence about the price and reference utility dependence of equivalence scales however, is fragmented and incomplete. The vast majority of the empirical studies that has employed subjective data as an information source about indifference curves in $\mathrm{q} - \mathbf{z}$ space impose overly simplistic functional forms on preferences that rules out the alleged price and base utility dependency of equivalence scales. Moreover, such functional forms typically predict behavioral responses that are inconsistent with empirical patterns of demand.

In this research we attempt to reconcile the empirical methodology commonly seen in the "satisfaction" literature with the technically more advanced literature on consumer demand. In doing so we propose to use both demand and data on subjective assessments of welfare in one integrated econometric framework. Subjective data is important because it is informative about indifference curves in $\mathrm{q} - \mathbf{z}$ space, and demand data is important because it is informative about indifference curves in q space conditional on $\mathbf{z}$. Because indifference curves in q space may be rather elaborate (which is why the literature on demand can be quite technical) we use variation in demands to obtain precise estimates of these parameters.

We estimate equivalence scales that are a little larger than the well-known modified OECD scales. Moreover, we find statistically significant price and base utility effects. Equivalence scales increase in base utility which means that rich household need more money to be compensated for the drop in welfare as a result of an increase in household size. This effect becomes smaller when the food/non-food price ratio increases.[32] Thirdly, we find that price effects are more important for low utility (i.e., poor) households, whereas the effects are practically non-existent for high utility (rich households). We find that for price effects two forces work in an opposite direction. For a fixed household income (or total consumption), a household responds optimally to an increase in household size by substituting expenditures towards food. When this household subsequently is compensated for the loss in welfare (to return to their original level of welfare), it will respond optimally by decreasing the expenditure share on food. The total price effect on equivalence scales depends on which one of the two effects is dominating. We find that for low utility (poor) household, the size effect is stronger than the compensation effect such that price increases increase equivalence scales. For high utility (rich) households, both mechanisms are in balance, such that equivalence scales are practically invariant to price changes.

To conclude. Whereas price and utility are statistically significant, one may argue that these effects are relatively small in qualitative terms. The benefit of using such elaborate scales in practice, may not outweigh the cost of gathering the appropriate data for constructing them.

---

[32] For extreme price regimes (price ratios must be more than 5 standard deviations larger than the mean) equivalence scales will decrease in utility.

Chapter 3

Subjective data in a test of the life cycle model

Abstract

The life cycle model is a cornerstone model in modern micro– and macro economics. The life cycle model offers a rationale for savings behavior, as individuals or households are assumed to value future utility in current period decisionmaking. A frequently employed test on the validity of the life cycle model is based on the orthogonality conditions implied by the first order condition of the problem: the Euler equation of consumption (Hall 1978). Such tests are convenient as they do not rely on a specification of income processes for example. Conventional approaches subsequently define preferences (i.e., the marginal rate of substitution between future and current consumption) and test the empirical validity of this –one specific– parameterization of the model. In this chapter I propose "self-reported changes in food consumption adequacy" as a direct proxy for the marginal rate of substitution of food consumption in two consecutive periods. Indonesian households were asked to rate the change in the adequacy of food consumption from a year ago until now on a five-point scale. The methodology has two clear advantages over conventional approaches. First, the test does not rely on ex ante specified preferences, such that it scrutinizes the validity of a life cycle model of unknown form (e.g., habit formation, within period nonseparabilities, etc. are implicitly allowed for).[1] *Secondly, conventional tests may suffer from low power as it can be difficult to statistically distinguish an "extended life cycle model" from alternative theories of behavior. In this chapter, I strongly reject the life cycle model with constant discount rates. The model with household specific discount rates is (only) borderline rejected.*

3.1 Introduction

In this research I present new empirical tests of the life cycle model. I test the predic-

---

[1] I assume that the household is the unit of decisionmaking, and life time utility can be written as an exponentially discounted sum of current and future "within period" utilities. Also, standard regularity conditions apply, such that within period utility/felicity is concave.

tions of the Euler equation of consumption without relying on arbitrary functional form assumptions on preferences. I use data on "recollected changes in consumption adequacy" as a directly observed proxy of the marginal rate of substitution between current and last year's consumption [see section 3.2.1].[2] I strongly reject the life cycle model with constant discount rates, and I (borderline) reject the life cycle model with household specific discount rates.

The approach I propose has two clear advantages over conventional approaches: first, the empirical validity of *a* life cycle model of unknown form is tested. This limits the set of possible alternatives associated with a rejection, and paves the way for estimating preference parameters once the model cannot be rejected. Second, whereas conventional approaches typically have difficulties to empirically disentangle "extensions" to the baseline life cycle model[3] and alternative theories of behavior, my approach should be immune to this. For example, there are many competing explanations for the –often observed– simultaneous increase in both consumption and income in the early phases of the life cycle.[4] A directly observed proxy for the marginal rate of substitution that I propose in this research can be helpful to discriminate between different explanations.

Ever since its introduction in the 1950's, the life cycle model is the dominant framework to think about savings behavior (Modigliani and Brumberg 1954). In its most general appearance, households value future (expected) utility when making current period decisions. The forward looking nature of the life cycle model is the main rationale for savings behavior. In anticipation of an income decrease, consumers respond optimally by rationing today's consumption and increase savings (Campbell 1987). Understanding savings behavior is important as (household) savings provide capital for economic growth. On the individual level, savings are an important means to insure consumption against adverse shocks in e.g., income.[5]

However, the life cycle model in and its applicability to real world decision-making has been criticized extensively in the literature (e.g., Carroll and Summers (1989) and Fernández-Villaverde and Krueger (2007), or see Browning and Lusardi (1996) for an overview "excess sensitivity" tests that are based on Euler equations). Different predictions of the life cycle model have been used to test the validity of the model. Carroll and Summers (1989) for example, study the cross sectional age-consumption profiles of fast and of slow growing economies. The life cycle model predicts that

---

[2] The data is drawn from the household consumption panel, which is part of a much larger data collection effort: the Indonesian National Socioeconomic Survey (SUSENAS) [see section 3.2 for more].

[3] An "extended life cycle model" allows for e.g., preference shifts due to changes in demographics and labor supply.

[4] e.g., liquidity constraints, nonseparability with demographics, habit formation, ad hoc (Keynesian) consumption.

[5] The precautionary motive is implicit in some versions of the the life cycle model (Deaton 1990).

"at a point in time the age-consumption cross-section profile should be less positively sloped in a rapidly growing country than in a slowly growing country. This is because in more rapidly growing countries the old are much lifetime-poorer than the young so consumption of the old will be much lower relative to consumption of the young (Carroll and Summers 1989)". They do not however find evidence for this prediction, as these profiles are similar across different countries.[6]

Often however, criticism on the life cycle model originates from statistical rejections of yet another, very much related prediction of the model. Hall (1978) showed that under some assumptions[7] "no variable apart from current consumption should be of any value in predicting future consumption". Many scholars have been challenging this prediction [see e.g., Browning and Lusardi (1996) for a summary of papers that have done this.]

Despite heavy criticism and frequent statistical rejections, the life cycle model still features a central role in economics. One may wonder why a model would survive all this critique? First, its basic principles have great intuitive appeal. It seems logical that individuals today, care about their future wellbeing in some way. Moreover, Deaton (2005) argues that "without it, we would have much less to say about many important issues, such as the private and public provision of social security, the effects of the stock market on the economy, the effects of demographic change on national saving, the role of saving in economic growth, and the determinants of national wealth".

A further reason for holding on to the life cycle model as the basic framework to think about savings behavior, is that empirical tests of the life cycle model often rely on strong *auxiliary* assumptions. Such assumptions are typically arbitrary (i.e., not implied by theory), such that a statistical rejection of the model can be interpreted as a simultaneous rejection of the basic principles of the model *and/or* of one or more of these auxiliary assumptions. That is, a rejection may be consistent with many alternative hypotheses, of which only one class is inconsistent with the life cycle model as a general representation of preferences: people do not smooth expected marginal utilities, but in fact do something else. Browning and Lusardi (1996) explain the survival of the life cycle model against many statistical rejections as follows: "the standard additive model has a number of *life belts* that may save it from drowning."

There are numerous examples of studies that argue that statistical rejections of the orthogonality conditions implied by the Euler equation are in fact due to errors in operationalizing the tests. Blundell et al. (1994) for example, find that excess

---

[6] Also, Fernández-Villaverde and Krueger (2007) and Alessie and De Ree (2009) show that consumption apparently tracks the hump shape in income over the life cycle. This is an apparent inconsistency with the theory as the hump in income can, at least in part, be anticipated.

[7] perfect capital markets, quadratic additive utility (certainty equivalence), rational expectations, and and exponential discounting where the interest rate equals the time discount factor.

sensitivity of consumption to anticipated changes in income can be explained by nonseparabilities with demographic variables (e.g., children). Carroll (2001) on the other hand, argues that log linearized Euler equations are inappropriate vehicles to estimate parameters (e.g., intertemporal substitution elasticities), particularly in a cross section of households. Also, Zeldes (1989) interprets excess sensitivity to income for low wealth households as evidence for the importance of binding liquidity constraints.

The fact that the model has great appeal *and* its ability to explain some real world phenomena, I argue, has motivated scholars to interpret a statistical rejection of the life cycle model as a failure of functional form (e.g., curvature of the utility function, non-separabilities within period or over time), institutional conditions (e.g., liquidity constraints), or estimation issues (small T dimension of survey data, or a failure of rational expectations). Indeed, not as a rejection of the fundamentals of the model. Within the tradition of testing the validity of the Euler equation of consumption, I relax some of these auxiliary assumptions (or life belts). In this research I use a directly observed proxy for the marginal rate of food consumption in two consecutive periods, and test the predictions of a general class of Euler equations of food consumption. The approach greatly narrows down the set of alternative hypotheses associated with a rejection of the model. As a result I propose a test on the validity of life cycle theory, rather than on the validity of *one specific* parametric representation of a life cycle model.

Conventional approaches typically employ log changes in consumption expenditures as a proxy for the marginal rate of substitution between this year's and next year's consumption. It is well known however, that such a linearization may be problematic as it eliminates the mechanism for precautionary savings [see e.g., Deaton (1991)]. Deaton (1991) considers that impatient households who cannot borrow typically engage in buffer-stock saving behavior. For these households, consumption growth is a crude, and for some purposes, inadequate proxy for the marginal rate of substitution of consumption in two consecutive periods.

The proxy that I propose is arguably much better. In this research I use recollected changes in consumption adequacy as a direct proxy of the marginal rate of substitution of consumption in two consecutive periods. A test on the basis of this proxy does not rely on functional form assumptions on preferences. That is, no assumptions need to be made on the curvature of the within period felicity functions, nor on within period separabilities. Also, the test is consistent under the usual types of intertemporal nonseparabilities (e.g., habit formation or durability). Finally, the approach is largely immune to measurement error that is known to plague consumption expenditure data obtained by household surveys (especially those where households do not use diaries, but are asked to remember what they have been consuming in the past). Especially, but not exclusively, in developing countries the problem of measurement error must be emphasized.

Although this approach solves a lot of problems, it also introduces new ones. Where data on direct measurements of preferences with subjective data is rapidly finding its way into mainstream economics, it is still largely unclear what many subjective measures actually mean, and, consequently, how they can be used. A priori, it is not obvious why recollected changes in (food) consumption adequacy is a good proxy for the marginal rate of substitution between current and lagged food consumption. Therefore, I will try to extensively justify my empirical strategy in section (3.2). Using subjective (or stated) preferences data as direct proxies of important economics concepts is relatively new, and I would argue, has great potential. I test the validity of the life cycle model with constant- and with household specific discount rates. I find a statistical rejection of both models. However, one could argue that the rejection of the model that allows for individual specific discount rates is not "extreme" (even if there is perhaps no obvious criterion by which one could value the severity of a rejection).

Finally, I test the validity of the life cycle model in the context of a developing country (i.e., Indonesia). An extensive body of literature in development economics focusses on risk sharing within communities [e.g., Townsend (1994)]. However, households can only "insure" idiosyncratic risk by sharing the burden within the community, so, even if households are perfectly able to diversify their idiosyncratic risk, it only contributes to eliminating the idiosyncratic risk in the total. In reality, the supposition of perfect risk sharing is too strict. This study is complementary to this literature as the uninsurable risk that the household is facing, should be dealt with in the same way as were risk sharing not possible.

3.2 Data description and the interpretation of *recollected changes in consumption adequacy*

For this research I use a new and unexploited household consumption panel data set drawn from Indonesian National Socioeconomic Survey (henceforth, SUSENAS). The original SUSENAS is a nationally representative survey among Indonesian households and interviews about 200,000 households on a yearly basis. Listings of consumption expenditure items of this broad survey however are rather limited. From 2002 up to 2004 the SUSENAS has sampled a subset (about 10,000 households) of the original SUSENAS to take part in a consumption panel.[8] In addition to a very detailed set of consumption expenditures and demographic variables, households were asked to rate

---

[8] The attrition rate of this panel data set is about 10% per year.

their respective food and non-food consumption as adequate or inadequate to satisfy their particular households needs (only the 2003 and 2004 wave). This data has been used by De Ree et al. (2009) to estimate the price and reference utility dependence of equivalence scales. The exact phrasing of the questions is as follows:

Q1. In the past month, has your food consumption been adequate for your household needs?[9] [10]

1. no
2. yes
3. more than adequate
4. do not know

This type of self-rated consumption adequacy data has already been part in various surveys across (mostly) developing countries [e.g. Pradhan and Ravallion (2000) for Jamaica and Nepal and De Ree et al. (2009) for Indonesia].

For a large part, this research relies on a question that has been asked as a follow-up to the question above. After households were asked to value the adequacy of consumption (a question on *levels*), households were subsequently asked a question about *change*:

Q2. How did the level of food consumption adequacy change, compared to last year?[11] [12]

1. deteriorated a lot
2. deteriorated a little
3. no change
4. improved a little
5. improved a lot

I will refer to this data as self-reported, or subjective recollections of changes in food consumption adequacy. Or in short, the recollections data $R = 1, 2, 3, 4, 5$. Figure (3.1) shows the frequencies of R in the data set (2003 and 2004 combined).

---

[9] The SUSENAS has the same question for nonfood. I do not consider this data for this research.

[10] In Bahasa Indonesian: Bagaimana tingkat pemenuhan kebutuhan konsumsi makanan rumah tangga selama sebulan yang lalu?

[11] The SUSENAS has the same question for nonfood.

[12] In Bahasa Indonesian: Bagaimana keadaannya jika dibandingkan dengan tahun lalu?

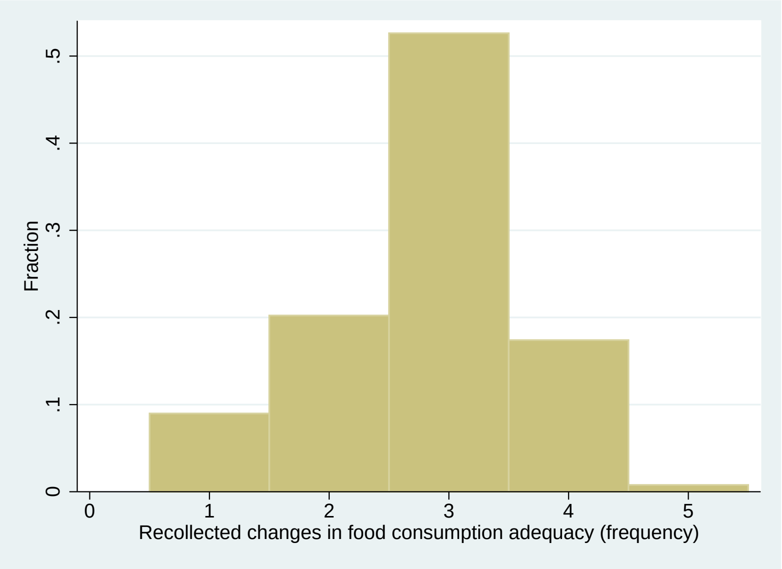


Figure 3.1: *Frequencies of households reporting different answers to the changes in food adequacy recollection question: from 1 (deteriorated a lot) to 5 (improved a lot).*

In what follows I will argue that the recollected changes in consumption adequacy is an appropriate proxy of the marginal rate of substitution of food consumption in two consecutive periods. I will consider degrees in *adequacy* of consumption and degrees in *satisfaction* with consumption to be the same thing. To my knowledge, direct measurement of this economic concept with stated or subjective preference data is a novelty. The empirical significance of this research depends however, to a large extent, on a correct interpretation of this data and it is consequently important to develop the intuition.

3.2.1 Interpreting recollected changes in consumption adequacy

The negative tension of *wanting* or *needing* a good or service can be resolved by satisfying the particular want or need (by e.g., acquiring the good or the service). Generally, satisfying a want generates a feeling of satisfaction with respect to that particular want. Satisfaction with x, therefore, is associated with relaxing tensions of wanting x. I am proposing therefore, in general, that the degree of *dis*satisfaction with x is proxying the negative *tensions of wanting* x. I have borrowed this interpretation of satisfaction from the psychological or psychoanalytical literature. Sullivan (1953) elaborates on the issue as follows:".. the relaxation of the tensions called out by lacks of this kind [lack of water, lack of oxygen, lack of sugar] I call *satisfaction* of the specific need which was concerned (p. 37)." Questioning of individuals about satisfaction therefore, is questioning individuals about the degree to which tensions of wanting are relaxed. Furthermore, and this is perhaps of additional importance to economists, Sullivan (1953) is directly linking tensions of wanting to behavior:".. a need, while it is in a broad biological sense disequilibrium, acquires its meaning from the actions or energy transformations which result in its satisfaction (p. 37-38)."

In economics, such tensions –that cause a feeling of dissatisfaction *and* spur action to alleviate them– are characterized by *marginal utilities.* The above logic therefore, suggests a link between *satisfaction* of wants or needs and *marginal utility.* As far as I am aware, this is a novel interpretation of satisfaction data that opposes (or perhaps refines) the widely adopted unification of (income) satisfaction and "utility". Where I find it convincing that satisfaction is related to alleviated tensions of wanting, it is perhaps not obvious why, in economics, these tensions are characterized by marginal utilities. Therefore I refer to early 19th century economics, the period when utility theory experienced some of their major breakthroughs. Early Austrian scholars, of which Gossen (1854, translated in English in 1983) was arguably the first, interpreted *marginal* utilities as a measure of the intensity or importance of the last want that is satisfied.[13] Marginal utility therefore is a measure of the psychological tension of wanting it.

Sullivan (1953) considers the possibility of a state of full "equilibrium" by satisfying all tensions ["lack of oxygen, lack of sugar, lack of water,..."]. In realistic economic situations however, individuals never seem to be able to relax all tensions, and consequently live in a permanent state of "disequilibrium" (in the sense that tensions of needs cannot be fully satisfied). Due to money, price, or other constraints, individuals balance out the different needs and wants they have given the constraints they face.

With regard to the life cycle model, the above logic suggests that a situation where the degree of satisfaction with real consumption in two periods corresponds one-to-one to a situation where marginal utilities of real consumption are equal in two subsequent periods. This prediction, moreover, resonates very well with casual remarks of e.g., Deaton (2005), who also seems to suggest that *need satisfaction* and *marginal utilities* are the same thing: "people...tailor their consumption patterns to their needs at different ages."

---

[13] A sidestep: On the basis of this logic the Austrians developed a theory about the shape of the utility function by supposing that the important wants (with high marginal utility) are satisfied first. Wants of lesser importance are satisfied later. They conclude from this idea that marginal utility is decreasing, and hence, utility is concave. Today, the utility functions we usually presuppose today are quasiconcave. However, the rationale for quasiconcavity is stems from the same logic.

It is however not to say that equalizing marginal utilities is primary objective of the consumer. Due to constraints on income, commodity prices, interest rates and time discount rates, the utility maximizing solution may not be that marginal utilities are equal over time, and nor do satisfactions with consumption.[14] Households may rationally accept a decrease in consumption satisfaction (i.e., an increase in marginal utilities) if one heavily discounts the future or if future commodity prices are high.

To further understand the working of this logic and its implications for testing the life cycle model I work out an example. I define a simple two-period life cycle model, where a household is supposed to maximize an additively separable utility function subject to a life time budget constraint. The within period felicity functions are allowed to depend on demographics z_t. I assume that borrowing or lending constraints are absent. The model can be written as follows:

$$\max_{\mathrm{c}_1,\mathrm{c}_2} E_1\left[u\left(\mathrm{c}_1,\mathrm{z}_1\right)+\beta u\left(\mathrm{c}_2,\mathrm{z}_2\right)\right] \tag{3.1}$$

where E_1 is the expectations operator conditional on information known to the decisionmaker at time 1. $u(\cdot)$ is a within period utility function that is strictly concave on the positive domain. c_t is real consumption at period t. Consumption in period 1 is the numeraire good and its price is normalized to 1. Consequently, the budget constraint can be written as:

$$\begin{aligned} \mathbf{x} &= \mathrm{c}_1+\frac{\mathrm{p}_2}{1+r_2}\mathrm{c}_2 \\ &= \mathrm{c}_1+\frac{1}{1+r_2^*}\mathrm{c}_2 \end{aligned} \tag{3.2}$$

z_t is a $\mathbf{x}$ is discounted life time income. p_2 is a price index capturing the commodity price change from period 1 to period 2. r_2 and r_2^* are the nominal and real rate of interest respectively. For expositional purposes I assume that both can be perfectly anticipated. Finally, suppose that marginal utility of consumption is *increasing* in household size, which means that there is a greater need for consumption if there are more mouths to feed. In this example I assume that $\mathrm{z}_1<\mathrm{z}_2$.

The utility maximizing solution to this problem is well-known and comes down

---

[14] Note that the life cycle model is a special case of a more general utility maximizing model. The life cycle model is unique in the sense that the "goods" that are consumed are effectively the same thing: real consumption, yet consumed at different points in time. The link between marginal utilities and satisfaction is not always so straightforward as it is in this case, and the establishment of a general relationship between satisfaction data and marginal utilities is beyond the scope of the current research. The issue here is that marginal utilities in a more general economic problem –such as the decision process over apples and oranges– are subject to normalizations (measuring goods in kilos or in ounces). "Consumption" in both periods, is logically normalized the same way in each periods.

to allocating life time income $\mathbf{x}$ such that the following condition holds:

$$E_1 \left[\frac{u_{c_2} \left(\mathrm{c}_2^*, \mathrm{z}_2 \right)}{u_{c_1} \left(\mathrm{c}_1^*, \mathrm{z}_1 \right)} \right] = \frac{1}{\beta \left(1 + r_2^* \right)} \tag{3.3}$$

where u_{c_t} is the partial derivative of u with respect to consumption at period t. c_1^* and c_2^* are period t decisions on consumption that maximize utility. Because $\mathrm{z}_1 < \mathrm{z}_2$ and because marginal utility of consumption is increasing in household size, note that if $\beta \left(1 + r_2^* \right) \approx 0$ the utility maximizing household typically consumes more in the second period than in the first: $\mathrm{c}_1^* < \mathrm{c}_2^*$.

But how would a household that lives in a world described by the model above answer the two questions that are posed by SUSENAS? Suppose that the uncertainty in life time income $\mathbf{x}$ has materialized in such a way that the household was *just* able to satisfy basic needs like food and shelter. Obviously, food and shelter is equally important in both periods such that marginal utilities of consumption have been equal in both periods. Note however, that satisfying those (basic) needs required more units of $\mathbf{x}$ in the second period as there are more mouths to feed (i.e., $\mathrm{z}_1 < \mathrm{z}_2$). As in both years, he *should* have answered affirmatively to the first question. Indeed, consumption is just adequate to satisfy the household's primary needs in the two periods of life.[15] In the second period of life, SUSENAS poses the additional question about how the "adequacy status" has changed with respect to a year ago. According to the above logic, the household head *should*[16] say, that consumption adequacy has not changed, because the household has managed to stay just adequate in both periods. Note that while adequacy stays fixed, real consumption expenditures has increased over time as a result an increase in household size.

If one accepts the recollected changes in food consumption adequacy as proxies of the marginal rate of substitution between today's and last year's food consumption, it offers great potential for testing the principles of the life cycle model. Using direct measurements of preferences in empirical operationalizations does not, to a large extent, depend on functional form assumptions that would otherwise frustrate empirical tests based on the Euler equation. Consequently, it narrows down the set of possible alternative hypotheses associated with a (potential) rejection. Moreover, but that beyond the scope of this study, it provides an additional source of information that can be used to study cross sectional and intertemporal heterogeneity of preferences, an issue that is often not considered.

---

[15] Pradhan and Ravallion (2000) show that the subjective poverty line based on this question largely coincides with the official poverty lines. This indicates that households, on average, consider themselves consumption adequate when they live above the official poverty lines

[16] The emphasis here is placed on the word *should* because in reality households may report something else.

3.3 Operationalization of the test: introducing a fairly general life cycle model

In the remainder of the text I assume that changes in recollected adequacy of food consumption is a proxy for the marginal rate of substitution (MRS) of food consumption in two consecutive periods. In this research I do not specify functional forms of the within period preference functions, nor do restrict preferences to be additively separable over time. Nevertheless, I impose a clear structure on preferences, to isolate a class of models that is usually referred to as life cycle models (some may have a different opinion about this). I suppose that at period t households maximize expected utility over the life cycle subject to a lifetime budget constraint. Life time utility is an exponentially discounted sum of current and future felicity functions. (I have dropped the household i subscript for expositional purposes.)

$$\max E_t \sum_{\tau=t}^{T} \beta^{\tau-t} u_\tau \left(c_\tau\right) \tag{3.4}$$

Subject to the life time budget:

$$\left(1+r_{t-1}\right) A_{t-1} + y_t + \frac{y_{t+1}}{1+r_{t+1}} + \frac{y_{t+2}}{\left(1+r_{t+1}\right)\left(1+r_{t+2}\right)} + \ldots =$$
$$c_t - \frac{c_{t+1}}{1+r_{t+1}} - \frac{c_{t+2}}{\left(1+r_{t+1}\right)\left(1+r_{t+2}\right)} - \ldots$$

E_t is the expectation operator, conditional on all information known to the decisionmaker at period t. E_t is short for $E\left[\cdot \middle| I_t\right]$, where I_t is the information set of the decisionmaker at period t. I use both ways of writing expectations in the chapter. u_τ is the within period utility (or felicity) function. u_τ is strictly concave in c_τ. The τ subscript on u_τ denotes its dependence on period τ (observed and unobserved) "taste shifters". These taste shifters may include demographic variables and choice variables such as durable consumption or labor supply variables, but also lagged values of c_τ itself, to allow for durability or habits in food consumption. c_τ is real food consumption at period τ. y_τ is real household income at period τ. r_τ is the real interest rate and β is the exponential discount rate.

For expositional purposes I assume that within period felicity is only function of current and one period lagged food consumption. Further generalizations are possible without disrupting the main idea of the study. Furthermore, and this is only to arrive at a convenient first order condition of the problem, I assume that the within period felicities depend on lagged consumption in a specific way.

$$u_\tau\left(c_\tau\right)=\tilde{u}_\tau\left(c_\tau-\alpha c_{\tau-1}\right) \tag{3.5}$$

where the subscript τ in $\tilde{u}_\tau\left(c_\tau\right)$ no longer includes lagged values of consumption. It is however, still permitted to to include demographic variables for example.

The first order condition of this problem becomes:

$$\begin{aligned} E_t\big[&\tilde{u}_t'\left(\cdot\right)-\left(1+r_{t+1}\right)\beta\tilde{u}_{t+1}'\left(\cdot\right)+\\ &-\alpha\beta\tilde{u}_{t+1}'\left(\cdot\right)+\alpha\left(1+r_{t+1}\right)\beta^2\tilde{u}_{t+2}'\left(\cdot\right)\big]=0 \end{aligned} \tag{3.6}$$

where $\tilde{u}_\tau'\left(\cdot\right)$ is the derivative of $\tilde{u}_\tau$ w.r.t. food consumption.[17] Hayashi (1985) shows that the above optimality condition can be greatly simplified if the real rate of interest r_τ is a constant r, and T (the final period of life) is far enough in the future (also used by Dynan (2000)) [see appendix B.2 for details].

$$E_t\left[MRS_{t+1}\right]=\frac{1}{\beta\left(1+r\right)} \tag{3.8}$$

where the marginal rate of substitution between period $t+1$ and period t consumption is defined as: $MRS_{t+1}=\frac{\tilde{u}_{t+1}'(\cdot)}{\tilde{u}_t'(\cdot)}$. Equation (3.8) can be rewritten as follows:

$$MRS_{t+1}=\frac{1}{\beta\left(1+r\right)}+\varepsilon_{t+1} \quad \text{where } E_t\varepsilon_{t+1}=0 \tag{3.9}$$

where ε_{t+1} is a forecast error.

Equation (3.8) is known as the Euler equation of consumption and implies testable restrictions on the data. Conventional approaches typically impose functional form assumptions on u_t' to estimate preference parameters and/or to test the empirical validity of that particular parameterization on the basis of consumption data. Not seldom hampered by data constraints, parametric assumptions on u_t' are highly restrictive. For example, the typical household survey exhibits limited information on the *stock* of durables. As a consequence, many authors are forced to assume that preferences over the service flows extracted from the stock of durables are separable from preferences over nondurable consumption. Such assumptions are often hard to justify and only because of practical considerations these kind of assumptions became generally accepted. Furthermore, assumptions are imposed on the curvature of the

---

[17]

$$\frac{\partial\tilde{u}_\tau\left(c_\tau-\alpha c_{\tau-1}\right)}{\partial c_\tau}=\tilde{u}_\tau'\left(c_\tau-\alpha c_{\tau-1}\right)\times\frac{\partial\left(c_\tau-\alpha c_{\tau-1}\right)}{\partial c_\tau}=\tilde{u}_\tau'\left(c_\tau-\alpha c_{\tau-1}\right) \tag{3.7}$$

within period preference function. A widely used functional form used in this context is the iso-elastic, or CRRA (constant relative risk aversion) utility function. Unlike the quadratic utility function, this functional form allows for a precautionary motive.

Although the CRRA utility function is arguably less restrictive than the quadratic one, its appeal is partly offset by problems in operationalizing the moment restrictions for estimating preference parameters and/or empirical testing.[18] In the CRRA case, for example, moment conditions implied by equation (3.8) are nonlinear in the parameters. Standard nonlinear GMM estimators that neglect the measurement error in consumption yield inconsistent parameter estimates (Amemiya 1985). Many authors therefore have resorted to (log) linear approximations of the Euler equation.[19]

In this research I do not impose such functional form assumptions, but rely on the recollection data as a direct proxy of the marginal rate of substitution in two consecutive periods (one year in this research). I assume that the recollection data relates to the MRS in the following way:[20]

$$\begin{aligned} R=1 &\quad\leftrightarrow\quad \alpha_3 < MRS \\ R=2 &\quad\leftrightarrow\quad \alpha_2 < MRS \leq \alpha_3 \\ R=3 &\quad\leftrightarrow\quad \alpha_1 < MRS \leq \alpha_2 \\ R=4 &\quad\leftrightarrow\quad \alpha_0 < MRS \leq \alpha_1 \\ R=5 &\quad\leftrightarrow\quad MRS \leq \alpha_0 \end{aligned} \tag{3.10}$$

The relationship between the MRS and the recollection data offers scope for testing the empirical tenability of the Euler equation under general functional form specifications on the marginal rate of substitution of food consumption in two consecutive periods. There are two clear advantages with this approach. First, my approach does not rely on (hard to defend) separability assumptions on preference (e.g., the usually imposed separability assumption between durable and nondurable consumption, or between consumption and leisure). A rejection of the subsequent empirical tests in section (3.4.1) should not be due to a misrepresentation of preferences.

Second, using a directly observed proxy for the MRS might also increase the power of Euler equation based tests. "Extended life cycle models", that allow for habits and demographics for example, may be hard to empirically distinguish from alternative theories of behavior. I consider once more the example I gave in the

---

[18] Also, the curvature of the CRRA is captured by a single parameter, hence introducing strong link between risk aversion and the intertemporal substitution elasticity.

[19] Alan et al. (2009) on the other hand, emphasize some of the problems with log linearizing the Euler equation and propose alternative estimation strategies, nevertheless relying on the assumption that measurement error is "classical" (multiplicative and independent of everything).

[20] Note that the answers to the recollection questions are inversely related to the marginal rate of substitution. That is, an improvement ($R = 4, 5$) relates to low marginal rates of substitution.

introduction. It is frequently observed that both consumption as well as income increase in the early phase of the life cycle. This is an apparent inconsistency with a baseline version of the life cycle model (i.e., a model that does not allow for demographics). It may subsequently be difficult to attribute this phenomenon to either changes in preferences due to demographic shifts (an "extended life cycle model") or to more add hoc (Keynesian) consumption rules. A direct proxy of the marginal rate of substitution however, should encapsulate the demographic extension to the baseline model. Consequently, such data can be used disentangle both explanations of the empirical phenomenon. That is, if the extended life cycle model is true, food consumption adequacy should not change while consumption increases along with income. If households are Keynesian consumers, adequacy with food consumption should rise with income.

3.3.1 Using the proxy in an empirical test of the life cycle model

There are a few advantages of using directly observed self-reported proxies of the MRS for an empirical test of the life cycle model. Nevertheless, the definition (3.10) already shows that R is only a categorical proxy of the marginal rate of substitution. This fact has some consequences that I discuss in this section.

The categorical nature of the answers to the recollection questions suggests using ordered probit (or logit) routines in some way. However, these routines rely heavily on distributional assumptions on the error terms in a latent model. In our case, the latent model underlying an ordered probit model would logically be a straightforward extension of equation (3.9). Hence, using ordered probit routines would place heavy distributional assumptions on the forecast error itself. This is problematic as life cycle theory does not have a lot to say about the conditional distribution function of the forecast error, apart from its conditional mean, i.e., $E_t \varepsilon_{t+1} = 0$. For example, it has been suggested that the conditional variance of the forecast error may be a function of on the level of cash-on-hand[21] for, so-called, buffer stock savers [see e.g., Carroll (2001)]. Furthermore, life cycle theory certainly does not claim that the forecast error is conditionally normally distributed.

The issue here is that an ordered probit model lays down a full parametric specification of the conditional expectation function $E_t R_{t+1}$. For testing the predictions of the life cycle model however, we can rely on less demanding parametric specifications. In other words, some of the assumptions underlying the ordered probit model might not be necessary, particularly, normality and the constancy of the higher order conditional moments of the forecast error ε_{it+1}.

---

[21] the sum of current income and assets

Two issues with the using the proxy call for further attention. First, the categorical nature of R, versus the continuous nature of the MRS. Second, the MRS is a cardinal concept (an MRS or two actually means something) whereas the R is ordinal.

The general question that I am interested in this section is the following: is it possible to come up with a moment condition on the basis the recollection variable R that can be used to test the validity of equation (3.8). To answer this question it is insightful to reformulate $E_t\left[MRS_{t+1}\right]$ without loss of generality (see appendix B.1):[22]

$$E_t\left[MRS_{t+1}\right] = \sum_j P\left(R_{t+1} = j \middle| I_t\right) E\left[MRS_{t+1} \middle| I_t, R_{t+1} = j\right] \qquad (3.11)$$

where $j = 0, 1, ..., J$ indicate the possible categories for R (which in my case are 5).

The difference between the expectation of the categorical variable R_{t+1} and the continuous MRS_{t+1} can be visualized as follows:

$$\begin{aligned} E_t[R_{t+1}] &= \sum_{j=1}^{J} j * P\left[R_{t+1} = j|I_t\right] && (3.12) \\ E_t[MRS_{t+1}] &= \sum_{j=1}^{J} E\left[MRS_{t+1} \middle| I_t, R_{t+1} = j\right] * P\left(R_{t+1} = j \middle| I_t\right) && (3.13) \end{aligned}$$

Furthermore, if the number of categories J goes to infinity, $E\left[MRS_{t+1} \middle| I_t, R_{t+1} = j\right]$ converges to a j specific constant: $\lim_{J\to\infty} E\left[MRS_{t+1} \middle| I_t, R_{t+1} = j\right] \to c_j$.[23] As a result, we can rewrite equation (3.12) and (3.13) as follows:

$$\begin{aligned} E_t[R_{t+1}] &= \sum_{j=1}^{J} j * P\left(R_{t+1} = j \middle| I_t\right) && (3.14) \\ E_t[MRS_{t+1}] &= \sum_{j=1}^{J} c_j * P\left(R_{t+1} = j \middle| I_t\right) && (3.15) \end{aligned}$$

As $c_1 < c_2 < ... < c_J$ there exists an monotonic mapping f such that $f(c_j) = j\ \forall j$.

---

[22] Note that $E_t\left[MRS_{t+1}\right] = E\left[MRS_{t+1} \middle| I_t\right]$, where I_t is the information set of the decisionmaker.

[23] In this research the number of categories is only $J = 5$, such that the subsequent derivations are approximate for the particular case at hand.

As a results:

$$
\begin{aligned}
E_t[R_{t+1}] &= \sum_{j=1}^{J} f\left(c_j\right) * P\left(R_{t+1} = j \middle| I_t\right) && (3.16)\\
&= E_t[f\left(MRS_{t+1}\right)] && (3.17)
\end{aligned}
$$

The monotonic transformation of the marginal rate of substitution $f(MRS_{t+1})$ in equation (3.17) can be rewritten as a Taylor expansion around a constant $\alpha = \frac{1}{\beta(1+r)}$.[24]

$$
\begin{aligned}
E_t[R_{t+1}] &= E_t[f(MRS_{t+1})] && (3.18)\\
&= E_t\big[f(\alpha) + f'(\alpha)\left(MRS_{t+1} - \alpha\right) + \\
&\qquad \frac{1}{2} f''\left(\alpha\right)\left(MRS_{t+1} - \alpha\right)^2 + \ldots\big] && (3.19)
\end{aligned}
$$

where under equation (3.8) we can rewrite to:

$$
E_t[R_t] = f(\alpha) + f'(\alpha) E_t \varepsilon_{t+1} + \frac{1}{2} f''\left(\alpha\right) E_t \varepsilon_{t+1}^2 + \ldots \qquad (3.20)
$$

This section so far has shown that that the conditional expectation $E_t R_{t+1}$ is a linear function of the conditional expectation of the forecast error $E_t \varepsilon_{t+1}$ *and* the conditional expectation of second and higher order conditional moments of the forecast error ε_{t+1}. Consequently, a statistical rejection of the validity of $E_t\left[R_{t+1} - \phi\right] = 0$ for example, where ϕ is a simple constant, does not necessarily reject the predictions of the life cycle model $E_t \varepsilon_{t+1} = 0$. Additional assumptions (or additional data[25]) are needed. In section 3.4 I present two series of tests. Both series are valid tests of the life cycle model under different assumptions.

3.4 Empirical section

So far, I have argued that the subjective recollected changes in food adequacy *should* proxy for the marginal rate of substitution of consumption in two subsequent periods. Yet, it does not mean that it actually does. Recently, many scholars recognize the added value of analyzing stated or subjective preferences data instead of, or in addition to revealed preferences data. Manski (2004) for example emphasizes the importance of measuring subjective expectations in surveys, to test or to relax the

---

[24] The choice of the constant is arbitrary. Yet, $\frac{1}{\beta(1+r)}$ is chosen for convenience.

[25] For example, the conditional variance of R is informative about the conditional variance of the forecast error. This avenue is left for future research.

usually (maintained) assumption on rational expectations for example.

The usefulness of stated preference data in applied work however, depends for a large part on whether the data actually measures what it is supposed to measure. Kapteyn (1994) for example tests whether stated preferences data measures "the same thing" as revealed preference data and finds some differences.[26] In section (3.2.1) I argue that the recollected changes in consumption adequacy *should* proxy for the marginal rate of substitution of consumption in two consecutive periods. In reality however, these variables might be very noisy, or perhaps even systematically biased. At the very least the recollected changes in adequacy should be positively correlated with changes in household food consumption. Note, for example, that in a boundary case, where the recollection variables are just measuring noise, the data will be consistent with the life cycle model, thereby *spuriously* supporting the life cycle hypothesis.

To "test" whether the recollected changes measure what they are supposed to measure I have regressed the real growth rate of food consumption on five dummies representing the five answers to the recollection question:

$$\begin{aligned}\Delta \ln c_{it} &= \zeta_0 + \zeta_1 \cdot D\left(R_{it}=1\right) + \zeta_2 \cdot D\left(R_{it}=2\right) + \zeta_4 \cdot D\left(R_{it}=4\right) + \\ &\quad \zeta_5 \cdot D\left(R_{it}=5\right) \qquad (3.21)\end{aligned}$$

Where the $D(R=j)$ are dummies that equal 1 if $R=j$. The baseline group is the group reporting no change. I have deflated all consumption variables with region specific food poverty lines.

In table (3.1) I estimate the parameters of equation (3.21) under four different assumptions on the error term. The first column are the baseline results from a simple OLS regression on equation (3.21). Three things can be noted here: first, the baseline group has a negative average growth rate of -0.07. Second, the group that reports $R=1$ (= deteriorated a lot) has a 3 percentage points lower growth rate than the baseline. $R=2$ cannot be statistically distinguished from the baseline. $R=4$ (slight improvement) on average reports a 7 percentage points higher growth rate that the baseline, and $R=5$ (improved a lot) reports a 13 percentage points larger improvement than the baseline. These results suggest that at least on average, the subjective questions measure what they are supposed to measure (the signs of the estimates are largely correct), and the estimated coefficients are of reasonable magnitude.

Third, the R^2 of the above regression is low. About one percent of the variation in the growth rate of food consumption can be explained by these dummies. The low R^2

---

[26] Also, he notes that his rejection may also be due to overly strict functional form assumptions in his empirical setup.

Table 3.1: *Regression results: equation (3.21)*

parameter	(1) pooled	(2) pooled	(3) FE	(4) FE
ζ_0	-0.076***	-0.076***	-0.073***	-0.073***
	(-17.55)	(-11.22)	(-10.83)	(-10.36)
ζ_1	-0.032***	-0.032*	-0.046**	-0.046
	(-2.798)	(-1.922)	(-2.006)	(-1.538)
ζ_2	0.002	0.002	0.004	0.004
	(0.205)	(0.154)	(0.236)	(0.195)
ζ_4	0.073***	0.073***	0.066***	0.066***
	(8.361)	(6.264)	(3.859)	(3.237)
ζ_5	0.131***	0.131***	0.102	0.102
	(3.310)	(2.810)	(1.249)	(1.145)
observations	17339	17339	17339	17339
R^2	0.006	0.006	0.003	0.003
F	25.19	12.14	5.404	3.535
# clusters		635		635
# groups			9619	9619

NOTE. Robust t-statistics reported in column 1 and 3. t-statistics based on clustered standard errors reported in column 2 and 4 (clustering on the village level). *** p<0.01, ** p<0.05, * p<0.1. Pooled OLS reported in column 1 and 2. Fixed effects at the household level reported in column 3 and 4.

(in combination with highly significant parameter estimates) indicate that there is a good deal of noise in either one of the two, or both measures of change. Survey data on consumer expenditures (both in developed and in developing countries) are known to be seriously infected with measurement error. Ahmed et al. (2007) estimate some properties of measurement error in recall expenditure data using the Canadian food expenditure survey. The data set combines recall and diary based expenditure data from the same households. By assuming that the diary based records are the correct ones, Ahmed et al. (2007) estimate that about "70 to 80 % of the cross sectional variation in recall consumption data is measurement error". A second problem is that their evidence suggests that measurement error is negatively correlated with the true values (i.e., measurement error is not classical).

Obviously, measurement error in consumption is exacerbated when applying first differences. In our data, the standard deviation of food consumption growth is 0.4, suggesting that more than 50% of the Indonesian households face year-to-year consumption drops or rises of more than 40%. These extreme statistics should raise eyebrows and suggests the importance of measurement error. The subjective recollections, on average, seem to reflect much more reasonable figures of consumption growth [see table 3.1]. The difference between the average household reporting a deterioration to an improvement is around 16% per year (which is the difference in the estimated coefficients). The apparent measurement error in consumption data is another important motivation to depart from tests based on observed consumption expenditures on the predictions of the life cycle model. If measurement error is a problem one needs to account for that in one way or the other, which often concerns making assumptions about the nature of the measurement error [see e.g., Alan et al. (2009) who assume that measurement errors are multiplicative, stationary and independent of everything. These assumption on the nature of the measurement error are at odds with empirical evidence (Ahmed et al. 2007)].

In column 2, 3 and 4 I challenge the results of column 1 by relaxing the i.i.d. assumption on the error terms. In column 2 I find roughly the same results after allowing for error dependencies within villages. Obviously, t statistics go down after this generalization. Column 3 and 4 report fixed effects (within) results on equation (3.21). Apart from the now insignificant ζ_5 parameter, the results do not change. Note, that the group that reports "improved a lot" is small [see figure 3.1]. After allowing for error dependencies within villages in column 4, the ζ_1 parameter is no longer significant. All in all, I find that the significance of the parameter estimates is not stable across columns. Yet, both the pooled results (column 1 and 2) and the fixed effects results (column 3 and 4) find a clear positive association between the recollections and changes in real food consumption. This pattern validates the empirical strategy of this research as both variables are linked, as they should be.

3.4.1 Testing the life cycle model

In this section I present two series of tests on the basis of the recollection questions R_{it+1}. (I reintroduce the i subscript in this section to denote households.) On the basis of equation (3.20) we can derive a general moment condition on the basis of R_{it+1} that relates to the moment condition implied by the life cycle model:

$$E_{it}\left[R_{it+1}\right] - \phi_{it} = f'\left(\alpha_i\right) E_{it}\left[\varepsilon_{it+1}\right] \tag{3.22}$$

where ϕ_{it} can be interpreted as an unknown parameter. ϕ_{it} is defined as follows:

$$\phi_{it} = f\left(\alpha_i\right) + f''\left(\alpha_i\right) E_{it}\varepsilon_{it+1}^2 + \dots \tag{3.23}$$

Testing the validity of $E_{it}\left[\varepsilon_{it+1}\right]=0$ therefore, is equivalent to testing the validity of:

$$E_{it}\left[R_{it+1}\right]-\phi_{it}=0 \tag{3.24}$$

The possibility that ϕ_{it} is i and t specific however, yields considerable problems for operationalizing a test on the validity of equation (3.24). We basically need additional assumptions, and/or additional data on ϕ_{it} to operationalize a test on equation (3.24) with the available data.[27] A general test on the validity of the Euler equation on the basis of the recollections data seems not possible.

In this section I choose to impose restrictions on ϕ_{it}. In the first series of tests (table (3.2) column 1 and 2) I assume that ϕ_{it} is constant across households and time, hence, I test the validity of:

$$E_{it}\left[R_{it+1}-\phi\right]=0 \tag{3.25}$$

A test on the basis of (3.25) places restrictions on the *type* of life cycle model that is tested (see 1. below).

1. $\alpha_i=\alpha$ needs to be constant across households and time. Because α is monotonically related to the discount rate, equation (3.25) relates to a life cycle model with constant discount rates.

In addition to the restrictions on the type of life cycle model that is tested, I also need some additional (auxiliary) assumptions on to validate the approach more generally:

2. $f\left(\alpha_i\right)\neq 0$. Otherwise the test has no power. This condition secures that there is a monotonic function that links R to the MRS.

3. The monotone transform f is linear OR higher order conditional moments of the forecast error are independent on the information set of the decisionmaker at time t, hence they are constant across households and time.

To summarize: if both 2. and 3. holds, testing the validity of $E_{it}\left[R_{it+1}-\phi\right]=0$ is a valid test of the life cycle model with constant discount rates.

This (restrictive) model with a fixed discount rate across households and time I subsequently refer to as the *pooled model.* Equation (3.25) implies the following orthogonality condition:

$$E\left[\left(R_{it+1}-\phi-E\left[R_{it+1}-\phi\right]\right)'\left(x_{it}-E\left[x_{it}\right]\right)\right]=0 \tag{3.26}$$

---

[27] In its most general appearance, where ϕ_{it} is an i and t specific unknown parameter, equation (3.24) cannot be generally rejected by the data.

Because ϕ is constant, the above can be simplified to:

$$E\left[\left(R_{it+1}-E\left[R_{it+1}\right]\right)'\left(x_{it}-E\left[x_{it}\right]\right)\right]=0 \tag{3.27}$$

where x_{it} is a subset of the decisionmaker's information set (i.e., $x_{it} \in I_{it}$). A logical test of the life cycle model with constant discount rates evaluates the sample equivalent of the population moment presented above, and test whether it is significantly different from zero. A way to evaluate this is by running a regression of R_{it+1} on a vector x_{it} and a constant c:

$$\textbf{pooled model:} \quad R_{it+1}=c+\gamma x_{it}+u_{it+1} \tag{3.28}$$

u_{it+1} is an error term. Estimation issues are discussed later on in this section.

In a next series of tests I allow the parameter ϕ_{it} to be a household specific, but constant across time:

$$E_{it}\left[R_{it+1}-\phi_i\right]=0 \tag{3.29}$$

Equation (3.23) indicates that this allows for considerably more flexibility, both of the *type* of model that is tested (i.e., household specific discount rates), but also on the restrictiveness of the auxiliary assumptions that validate the approach more generally (e.g., higher order conditional moments may be household specific constants, rather than constant across households and time). Allowing for a household specific parameter ϕ_i no longer requires discount rates to be constant across households, and hence allows for a broader class of models under the null hypothesis.

Moreover, it also allows for the possibility that households make systematic, household specific errors in answering the recollection questions. This is important as respondents may not answer the recollection question in the same way as (I suppose) they should. Where the recollections questions should reflect a *change* in food consumption adequacy, the variable might be contaminated with household specific moods. The interviewer may have interviewed the respondent on a bad or a good day (increasing the general noise levels in such measures). However, day to day changes in moods should not be correlated with important elements of the household information set, like household size, or household consumption. Consequently, it should not cause trouble in the subsequent tests. What *is* important for interpreting the outcomes is that wealthier households might be in a better mood in general, and therefore more likely to report a positive answer to any (mood related) question. Allowing for a household specific parameter ϕ_i accounts for this possible impurity in the answers to the recollection questions.

A test on the basis of (3.29) places restrictions on the *type* of life cycle model that is tested (see 3. below).

4. α_i is a household specific parameter. Because α_i is monotonically related to the discount rate, equation (3.29) relates to a life cycle model with household specific discount rates that are constant over time.

In addition, the requirements for the additional (auxiliary) assumptions to validate the approach more generally are somewhat weaker than in 2. and 3.

5. $f(\alpha_i) \neq 0$. Otherwise the test has no power. This condition secures that there is a monotonic function that links R to the MRS.

6. The monotone transform f is linear OR higher order conditional moments of the forecast error are allowed to be household specific constant. Still, they are constant over time.[28]

To summarize: if both 5. and 6. holds, testing the validity of $E_{it}\left[R_{it+1} - \phi_i\right] = 0$ is a valid test of the life cycle model with household specific discount rates.

I subsequently refer to the model with a household specific discount rates to as the *first difference model*. Equation (3.29) implies the following orthogonality condition:

$$E\left[\left(R_{it+1} - E\left[R_{it+1}\right]\right)'\left(x_{it} - E\left[x_{it}\right]\right)\right] - E\left[\left(\phi_i - E\left[\phi_i\right]\right)'\left(x_{it} - E\left[x_{it}\right]\right)\right] = 0 \quad (3.30)$$

The life cycle model therefore, under the assumption of household specific constants, does not predict that R_{it+1} and x_{it} are uncorrelated. Instead, it predicts that the covariance of R_{it+1} with x_{it} is equal to the covariance of ϕ_i with x_{it}. The latter covariance term is generally non-zero. A sample equivalent of the above condition however, cannot be straightforwardly evaluated as the ϕ_i's are unobserved. The ϕ_i's need to be controlled for in some way.

The ϕ_i's can be differenced out from the baseline prediction of the life cycle model. Theory predicts $E[R_{it+1} - \phi_i | I_{it}] = 0$, but also $E[R_{it} - \phi_i | I_{it-1}] = 0$ logically. Subtracting the one from the other yields:

$$E[R_{it+1} - \phi_i | I_{it}] - E[R_{it} - \phi_i | I_{it-1}] = E[R_{it+1} | I_{it}] - E[R_{it} | I_{it-1}] = 0 \qquad (3.31)$$

which is a "new" moment condition implied by the theory, yet, this one does not depend on the unknown household specific effects. By applying the law of iterated expectations we can define a new (testable) condition:

$$E[R_{it+1} - R_{it} | I_{it-1}] = E[\Delta R_{it+1} | I_{it-1}] = 0 \qquad (3.32)$$

---

[28] On could argue that this is problematic as the conditional variance of the forecast error is time varying for buffer-stock consumers.

The condition implies again:

$$E[(\Delta R_{it+1} - E\,[\Delta R_{it+1}])'(x_{it-1} - E\,[x_{it-1}])] = 0 \tag{3.33}$$

Again, a logical test of the life cycle model with household specific discount rates is a test on the statistical significance of the sample equivalent of the population moment presented above. I evaluate this by running a regression of ΔR_{it+1} on a vector x_{it-1} and a constant $\tilde{c}$:

$$\textbf{first difference model: } \Delta R_{it+1} = \tilde{c} + \tilde{\gamma} x_{it-1} + \tilde{u}_{it+1} \tag{3.34}$$

In the subsequent sections I present the results of the pooled and the first difference model.

Before moving on to results of the tests, I wish to point to two important caveats. Both issues relate to the issue of approximating population moments with the associated sample equivalents. These issues have been raised before and are not specific to my approach. Nevertheless, they are important. The first issue is that we need to rely on *rational expectations* in order to secure that sample moments approach population moments in large samples [see e.g., Manski (2004) and Kapteyn et al. (2009) who try to relax the assumption on rational expectations using data on subjective expectations].

A second consistency issue has been raised by Chamberlain (1984). Forecast errors, average out to zero if $T \to \infty$, because of the serial independence that is implied by equation the Euler equation of consumption. Instead, Chamberlain (1984) notes that there may be "common components" in the forecast error as a result of aggregate shocks (to income for example). Consequently, forecast errors do not necessarily average out to zero if the cross sectional dimension $N \to \infty$. This is particularly problematic if those common components are correlated with the some variables in the household's information set in a cross section of households. This would be the case if, for example, wealthier households benefit more from unanticipated events (i.e., shocks) to the global economy than the poor.

Chamberlain (1984) was the first to note that where sample moments are expected to converge to the population moments if T goes to infinity, it may not converge when N, the number of households in the sample, goes to infinity. In the subsequent empirical operationalizations I allow for region specific time dummies in an attempt to filter out (at least) part of the common components in the forecast error (this approach is suboptimal as Altug and Miller (1990) for example, show that aggregate shocks are only effectively accounted for under the assumption of complete markets). In the remainder of the study I rely on N asymptotics in the sense that I assume that sample moments converge to population moments when N goes to infinity. However, with interpreting the outcomes of the tests we need to keep this possible consistency problem in mind.

Empirical tests of the life cycle model

I present the results of the OLS regression on the pooled model (equation (3.28)) and the first difference model (equation (3.34)) in table 3.2. As regressors I include the logarithm of real household consumption expenditures, and the logarithm of household size. Both variables are evaluated one period (a year in this case) before the household reported R. Further, I have included interactions of province and time dummies in an attempt to filter out aggregate components in the forecast error and to control for region specific interest rates (i.e., households in the data came from 30 Indonesian provinces[29]). The substantive results do not depend on including or excluding these dummies. Furthermore, I allow for error dependence at the village level. It is not unlikely, that unanticipated events are correlated within villages (or cities). Neglecting this correlation would lead to an overestimation of the t statistics, which might cause a spurious rejection of the life cycle model. The frequencies shown in figure (3.1) can be used as a baseline to interpret the marginal effects presented in column 1-3 of table (3.2).

Column 1 and 2 present the results of the pooled model. Both consumption and household size are (highly) significantly correlated with changes in food consumption adequacy. All else equal, wealthier households are more likely to report a positive change in adequacy. Equation (3.25) predicts however that wealthier households should not be significantly more likely to experience an improvement in consumption adequacy over the subsequent year. Column 3 is a validity check of the empirical approach and tests whether consumption growth is correlated with changes in adequacy (as it should).[30]

Column 4 and 5 are the results of the first difference model. Initial income is no longer significant in these regressions, a result that is consistent with the predictions of a life cycle model that allows for heterogeneity in the discount factor. On the other hand, I do find that changes in the recollection questions are positively (partially) correlated with household size. The rejection of the model with household specific discount rates is no longer related to income, yet it seems somehow related to household size. Again, column 6 reports the validity check on the empirical approach in general. The fact that the logarithm of period $t-2$ consumption does not correlate with ΔR_{it+1} in column 4 and 5, may be due to the possibility that ΔR_{it+1} just measures noise. The significant correlation between ΔR_{it+1} and

---

[29] There are 30 provinces, 286 districts, 636 subdistricts, 659 villages in the data.

[30] related to the issue described in section 3.4.

Table 3.2: *Results: pooled + first difference*

VARIABLES	(1) pooled	(2) pooled	(3) pooled	(4) first diff.	(5) first diff.	(6) first diff.
ln real tc $t-1$	0.194***	0.262***				
	(10.22)	(10.61)				
ln h.h. size $t-1$		-0.145***				
		(-6.461)				
Δln real fc			0.129***			
			(5.381)			
ln real tc $t-2$				0.008	-0.037	
				(0.305)	(-1.072)	
ln h.h. size $t-2$					0.096**	
					(2.570)	
Δln real fc $-$ Δln real fc $t-1$						0.091*** (3.665)
constant	2.690***	2.760***	2.810***	0.129	0.084	0.195***
	(27.81)	(28.71)	(156.4)	(1.256)	(0.824)	(8.828)
Observations	17340	17340	17339	7721	7721	7720
region×time dummies	yes	yes	no	yes	yes	no
R^2	0.066	0.070	0.004	0.016	0.017	0.003
F	6.983	7.038	28.95	1.656	1.825	13.43
# clusters	635	635	635	602	602	602

Note. t-statistics (in parentheses) are calculated on the basis of clustered standard errors. Standard errors are clustered at the village level (*** p<0.01, ** p<0.05, * p<0.10). Δ denotes the first difference transformation, such that $\Delta x_t = x_t - x_{t-1}$. Dependent variable: recollected change in food consumption adequacy at t: in levels (col 1-3) and first differences (col 4-6).

Δln real $\text{fc}_{it+1} - \Delta ln$ real fc_{it} I report in column 6 indicates that it does not.[31]

At face value the results of table (3.2) should be interpreted as a rejection of the life cycle model (both the one with constant discount rates and the one with individual specific discount rates). I have not specified a specific alternative hypothesis such

---

[31] Including age, age squared, or dummies for different age groups, did not significantly improved the fit of the model and are omitted.

that the cause of the rejection is not a priori clear. In fact, the rejection may be due to a violation of one or more of the auxiliary assumptions that I have taken to arrive at the final results. In the remainder of the section I will discuss three of the potentially important auxiliary assumptions one by one. Subsequently, I discuss a fourth alternative that we are typically most interested in: a possible alternative theory of behavior. My test does not straightforwardly discriminate between the four options. However, the test greatly limits the set of alternatives associated with a rejection. By using the recollections data as a direct proxy for the marginal rate of substitution, there is no need to look for an explanation in nonseparabilities (with leisure or durables), habit formation, measurement error, or curvature issues (e.g., precautionary savings).

1. The proxy problem. The recollection data is a proxy of the marginal rate of substitution, but it is not the same [see section (3.3.1) for a discussion]. The reason for writing this chapter is that subjective data in general, and the recollection data in particular, can be used to relax some (often hard to defend) functional form assumptions on preference needed to operationalize a test on the theory. Section (3.3.1) however, has also identified a few problems with using the recollection data to proxy the marginal rate of substitution.

There are two obvious explanations for the positive correlation between consumption and R_{it+1} in column 1 and 2. First, wealthier household discount the future less strongly than the poor, and hence, expect adequacy of food consumption to increase at a higher rate. Second, wealthier households are more likely to *report* an improvement in food consumption adequacy whereas it merely reflects the notion that wealthy households are more likely to report something positive more generally. A comparison of the results of column 1 and 2 versus column 4 and 5 seem to suggest that wealthier households generally are more likely to report improvements. In this research we cannot distinguish between the two competing explanations. Nevertheless, the first difference approach controls for both and produces, arguably, a more appropriate test on the fundamentals of the model.

Section (3.3.1) identifies a number of different reasons of why the proxy does not work well in an empirical test of the life cycle model. If the transformation f is nonlinear, the consistency of the test on the basis of the proxy variable is only a valid test of the life cycle hypothesis if higher order moments of the forecast error are constant, in column 1 and 2, or household specific constants, in column 4 and 5. This may be problematic. It has been suggested that the conditional variance of the forecast error is a function of cash-on-hand (or income) for buffer-stock consumers. Buffer-stock consumers are typically inclined to consume more than current income (i.e., they are "impatient"). However, the prospect of a drop in consumption due to a binding liquidity constraint in the future, prevents them from doing so. Buffer-stock consumers therefore balance between their tendency of consuming more than current income (when cash-on-hand is high, and the risk of binding liquidity constraint in the near future is low) and their tendency to save to insure themselves against extreme drops in consumption (when cash-on-hand is low, and the risk of binding liquidity constraint in the near future is high). One can show that this mechanism produces a dependency of the conditional *variance* of the forecast error on cash-on-hand. If total consumption can be taken as a proxy for cash-on-hand, I do not find evidence for buffer-stock behavior in column 4 and 5. The fact that the tests reported in column 4 and 5 allow for household specific differences in the higher order conditional moments of the forecast error, makes is far less clear why the conditional variance of the forecast error would be a function of household size, simply because household size is relatively constant over time.

2. The small T, large N problem. When there are aggregate, common components in the forecast error that are not adequately controlled for by including region-time dummies, the results of table (3.2) are inconsistent under the null [see section (3.4.1) for a discussion]. Common components of the forecast error may be correlated with initial consumption in small T samples. One could think of a income shock to the global economy, such that (permanent) income of all households increases. If large households respond differently to this shock than small households, we could find correlation between household size and the recollected changes in adequacy in small T panels, even if the life cycle model is correct. This is a reasonable explanation of the results that I cannot rule out.

3. The rational expectations assumption. Only under rational expectations the sample moments converge to population moments in large samples. The assumption of rational expectations is a strong one however. It is not unlikely that households make systematic errors when maximizing expected life time utility. As a result, their *subjective* expectation of future states of the world (say their future income or of their own future preferences), do not match their empirical counterparts on average. The significant parameter estimate with household size in column 5 may reflect a failure of rational expectations. Again, if a failure of the rational expectations assumption fully accounts for the significant parameter in column 5, the life cycle model as such, should not be disqualified.

4. An alternative hypothesis of behavior. Households may not attempt to maximize expected (exponentially discounted) utility over the life cycle, but instead do something else. Households may behave more add hoc and simply adjust their consumption to current income. Also, households may be boundedly rational and maximize *some* objective function, but not the one that would maximize expected utility over the life cycle (e.g., myopic habit formation, or some naive version of hyperbolic discounting).

Many authors have been more explicit about the alternative hypothesis that is associated with a rejection of the life cycle model. The large literature on so-called "excess sensitivity" tests for example also rely on the predictions of the Euler equation, but generally have a more specific alternative hypothesis in mind. The Euler equation of consumption implies that consumption growth is uncorrelated with anticipated changes to income. Nevertheless, many empirical studies have found that that this correlation is positive and significant, thereby finding that consumption is "excessively sensitive" to income [see e.g., Browning and Lusardi (1996) for an overview]. Excess sensitivity however can be straightforwardly explained by binding liquidity constraints (for some households), or by simple add hoc consumption behavior where simply a fixed fraction of current income is consumed ("Keynesian" consumption). Both phenomena generate excess sensitivity of consumption to income.

Interpreting my results as evidence for excess sensitivity depends on a correlation between household size and anticipated changes in income. I find this unlikely. In contrast however, the fact that initial consumption is insignificant in column 4 and 5 may be even interpreted as evidence *against* excess sensitivity due to liquidity constraints. Households with binding liquidity constraints expect, almost by definition, an income increase (and a associated decrease in marginal utilities) and should be more likely to report an improvement in consumption adequacy a year later when the income increase has materialized (in expectation). Zeldes (1989) for example, argues that low wealth households are more likely to be liquidity constraint and indeed finds evidence for that hypothesis (by finding excess sensitivity of consumption growth to initial income for low wealth households only).[32] If low consumption households are indeed more likely to be liquidity constrained, they should be consequently more likely to report an improvement in adequacy. My findings therefore do not endorse the importance of liquidity constraints for low income households.

To conclude this section I would like to mention that the life cycle model with household specific discount rates, although statistically rejected, does not seem to be so bad. Both consumption levels and age variables (not reported in table (3.2), but available upon request) are insignificant in column 5. The household size variable is (borderline) significant with a t statistic of around 2.5. Even though the t statistics are calculated on the basis of clustered standard errors at the village level, there may be other dependencies in the error terms that are not accounted for. t statistics of around 2.5, therefore, should be considered an upper bound on some other, perhaps more appropriate t statistic that takes this dependencies into account. Consequently, t statistics of around 2.5 might not be overly convincing.

---

[32] Zeldes (1989) argues that income, in a group of low wealth households, is correlated with the Kuhn Tucker multiplier associated with the borrowing constraint. In high wealth groups, he argues, households are not liquidity constraint, hence, Kuhn Tucker multipliers are consequently zero.

Also, I would like to point to the low R^2 in all table 3.2 regressions. If the life cycle model is true, the fit indeed should be low. However, the low fit can be misleading evidence in favor of the life cycle model due to measurement error in both l.h.s. and r.h.s. variables. Note that the validity check regressions in column 3 and column 6 also report low R^2's, supporting the measurement error hypothesis.

3.5 Conclusion

In this chapter I present a new empirical test of the life cycle model. I test the orthogonality conditions implied by the Euler equation of consumption, that is, any variable available to the decisionmaker at period t, should be uncorrelated with the marginal rate of substitution between period $t+1$ and period t (food) consumption. This research departs from using log differences in food consumption as a proxy for this marginal rate of substitution (i.e., log linearized Euler equation), and uses "qualitative recollections of changes in food consumption adequacy" as a directly observed proxy for the marginal rate of substitution between current and last year's food consumption. The test consequently does not rely on auxiliary (and consequently arbitrary) functional form assumptions on preferences and therefore proposes a much "cleaner" test on the fundamentals of the model. I test the validity of the life cycle model that allows for habit formation (or intertemporal nonseparability more generally) and within period preference functions of unknown form (thereby allowing for nonseparabilities with durables or labor supply variables).

I use data from a consumption panel of Indonesian households drawn from the Indonesian National Socioeconomic Survey (2002-2004). Apart from the standard content of consumption surveys, households were asked to rate their household's food consumption as adequate or inadequate for their household's needs. In a followup question households were asked to rate the change in food consumption adequacy with respect to their adequacy levels a year ago, on a five point scale. In section 3.2.1 I argue that this variable is a proxy of the marginal rate of substitution of current over last years food consumption. Note that this proxy is only partial as the recollections variable is a categorical variable, whereas the marginal rate of substitution is continuous.

I strongly reject the life cycle model with constant discount rates. Wealthier households are more likely to report a increase in adequacy a year later. This result however, is also consistent with a general misinterpretation of the recollection data. Wealthier households, for example, may be generally more likely to report a positive answer to any "satisfaction" related question. In a second series of tests I simultaneously allow for this "reporting bias" as well as for household specific discount rates.[33]

---

[33] With reporting bias I mean that household's capability to judge specific is affected by moods

I subsequently (borderline) reject the life cycle model with household specific discount rates. The rejection of this version of the life cycle model is due to a significant correlation of the recollection data in first differences, with household size. Initial consumption and age variables are not significantly correlated with recollection data in first differences.

Subjective data, or stated preferences data is a interesting source of information that has great potential for relaxing some of the commonly maintained assumptions in economics (e.g., rational expectations, constant preferences, etc.). Interpreting stated preferences data in relation to economic concepts however, I find, is yet in its infancy. Income satisfaction data for example, is often rather crudely taken as a measure of "utility". Much progress can be attained by studying the information content of stated preferences data. Theories (with clear predictions) need to be developed about the relationship between different types stated preferences data and on the relationship between revealed preferences data and stated preferences data. A black-and-white distinction between *experienced* utility and the utility concept that is used in demand analysis (*decision* utility), I find largely unsatisfactory [see Kahneman et al. (1997) for more on this]. In this research I try to contribute to studying this link.

I have relaxed the functional form requirements usually imposed when testing the validity of the life cycle model (such that only one specific life cycle model is tested, not the model more generally). Interesting progress can be attained by thinking further about appropriate questions that relate to the intertemporal maximization problem. An interesting combination of questions could be the following: 1. How do people expect their satisfaction or adequacy with food consumption to develop over the coming year(s). 2. How do people expect their income to develop over the coming year(s). Such data would allow for a clean test on the basic principles of the life cycle model with a mere cross section of households.

---

in general.

Published as: Rob Alessie and Joppe de Ree – *"Explaining The Hump In Life Cycle Consumption profiles,"* De Economist, vol. 157, no. 1, pp. 107–120, March 2009.

Chapter 4

Explaining the hump in life cycle consumption profiles

Abstract

This chapter documents life cycle (or age) profiles of (log) household income, durable and non-durable consumption for Dutch households after explicitly controlling for time (or business cycle) effects and birth cohort effects. We find that both measures of consumption as well as income is clearly hump shaped over the life cycle. Hence, real consumption per household seems to track income over the life cycle. This empirical regularity is hard to reconcile with basic specifications of the life cycle model. We further document life cycle profiles of demographic and labor supply variables. We argue that part, but not all, of the hump in consumption may be explained by household composition variables. Durable consumption per adult equivalent stays approximately flat until age 60 after which it drops dramatically. This phenomenon may be partly explained by a decrease in work related durable expenditures after retirement. Non-durable consumption per equivalent adult increases steadily until age 55 and stays approximately flat after that.

4.1 Introduction

This chapter examines the life cycle profiles of durable and non-durable consumption expenditures using a series of repeated cross sections drawn from the Dutch consumer expenditure survey [source: Statistics Netherlands]. The relevance of our study can be motivated in at least three ways.

Firstly, the consumption profiles provides prima facie evidence for the validity of the baseline life cycle model, the general framework for analyzing household consumer behavior over the life cycle. This model assumes that the agents maximize intertemporally additive expected utility, subject to a lifetime budget constraint. Moreover, the model assumes perfect capital markets (no liquidity constraints), complete markets and within period utility functions that are additively separable between non-durables, durable service flows, leisure and demographics. A key prediction of this version of the life cycle model is that changes in (the marginal utility of) non-durable consumption are uncorrelated with predictable changes in income. Like other studies (see e.g. Fernández-Villaverde and Krueger (2007)) we find however, that consumption seems to *track* income. An empirical regularity that is hard to reconcile with these basic specifications of the life cycle model.

Similarly, empirical studies using UK and US data typically find that both non-durable consumption and income drops at retirement (see e.g. Banks et al. (1998) and Bernheim et al. (2001)). One could be tempted to interpret these findings as evidence for the hypothesis that people have not saved enough for retirement. However, such policy relevant conclusions should be drawn with great caution, simply because these are based on a very simple version of the life cycle model.

Secondly, we assess the relevance of possible extensions to the basic life cycle model. The fact that consumption seems to track income over the life cycle may be explained by non-separabilities in the utility function. We report for example, that household size is also hump shaped over the life cycle and can account for a part of the hump shape in consumption.

Thirdly, the age-consumption profiles are useful empirical benchmarks for comparing the life cycle predictions of macro simulation models such as the one constructed by Auerbach and Kotlikoff (1987) and the GAMMA model of the Netherlands Bureau for Economic Policy Analysis (CPB) (see Draper and Amstrong (2007)). These models typically abstract from business cycle–, cohort– and demographic effects. In our empirical analysis we explicitly control for these effects. Predictions from these macro economic models may therefore be readily compared with their empirical counterparts that we document in this chapter.

For this research we borrow heavily from Fernández-Villaverde and Krueger (2007) who analyze the US consumer expenditure survey. This allows to compare our findings based on Dutch data with those of Fernández-Villaverde and Krueger (2007). Such a comparison is interesting since Dutch households face an institutional environment (e.g. capital markets and pension systems) that is quite different from the American one. Like Fernández-Villaverde and Krueger (2007), we find that after controlling for cohort and time effects, both durable and non-durable consumption show a clear hump shape over the life cycle. Only part of the hump can be explained by changes in household composition. Moreover, the data suggest that durable consumption is at least partly work related as durable consumption drops dramatically around retirement.

This chapter is organized as follows. First we describe the data we use in section (4.2). In section (4.3) we provide some details on the econometric methodology. Especially, we pay attention to our method to disentangle age, period and cohort effects. Moreover, we document the consumption and income profiles as functions of age and birth year respectively. Finally, we study the effects of controlling for family size. In section (4.4) we discuss possible additional explanations for explaining that part of the hump that cannot be attributed to a change in household size. Finally we summarize and conclude in section (4.5).

4.2 The data

The data used for this study are drawn from the Dutch budget survey (Budgetonderzoek) which is held by Statistics Netherlands at an annual basis. We use 23 waves that cover the period 1978-2000. The survey collects data on 2000 to 3000 individual households per wave. Only in the year 1991, the budget survey has been conducted among about 1000 households. The survey relates expenditures on a very detailed set of consumption categories to information on household composition and income. Furthermore, the survey contains information on income, family composition and background information on all members of the household (age, education etc.). In addition the data contains information on whether the household head, and the partner work fulltime, part time or not at all.[1]

The budget survey is not entirely representative for the Dutch population of households. One of the reasons for this is that every 5 years Statistics Netherlands constructs a weighting scheme from the budget survey which is used for price index calculations of the employed. To this end households where the head of the household is employed are oversampled in 1980, 1985, 1990, 1995 and 2000. Moreover, other types of households are over-represented in some other waves [see Kalwij et al. (1998) for more details]. Therefore, we use in our analysis the sample weights provided by Statistics Netherlands. Some sample selections have been applied. Firstly, we have excluded the households whose head is younger than 21. Secondly, those households with heads born before 1906 or after 1970 are also removed from the sample. Thirdly, the expenditure data of a few households showed some serious inconsistencies and are also removed from the sample.

Nondurable consumption is constructed as the sum of the expenditures on all non-durable items. Non-durable items are defined to depreciate within a year. They include among other things expenditures on food, clothing and rent (or imputed rents for house owners). It is reasonable to assume that non-durable *expenditures* and non-durable *consumption* are approximately equal. For durable consumption this is different. By definition, durable consumption items do not depreciate within a year. Durable expenditures therefore do not equal durable consumption services. Households are assumed to derive utility from the service *flow* of the durable stock.

---

[1] For the 1988-1991 waves the data also contains information on actual hours of labor supply by both the head and the partner. The dummies however can explain approximately 80% of the variation in hours of labor. See De Ree and Alessie (2008) for more details.

The flow however, is typically not recorded in expenditure surveys. It is important to keep this in mind when analyzing the life cycle patterns of durable consumption *expenditures*. Durable expenditures include expenditures on cars, furniture, but also investments in schooling.

The Dutch consumer expenditure survey is unique in the sense that it contains for home owners direct data on the rental value of their dwellings.[2] In other words, contrary to Fernández-Villaverde and Krueger (2007) we observe the service flow from the durable stock 'housing'. We have added this service flow to our measure of nondurable consumption. Utility theory suggests that this is a reasonable procedure. We also performed a sensitivity analysis with a nondurable consumption measure that excludes both the rent and the rental value.

The income and expenditure series are expressed in 1978 prices using the consumer price indices published by Statistics Netherlands [source: `http://statline.cbs.nl/`].

4.3 Consumption over the life cycle

At the minimum, one has to control cohort and calendar year effects if one estimates life cycle profiles of (non)durable consumption and income. People coming of age in different times have different preferences towards e.g. risk. Generations who endured the Great Depression in the 1930's might be more thrifty or more risk averse than other cohorts. An alternative view is that whereas preferences may be identical across cohorts, the economic conditions of the past are very different from the present. These considerations lead to the supposition of cohort or generation effects. Furthermore, calender year or business cycle effects might also seriously distort the cross-sectional life cycle profiles of consumption. In the subsequent econometric analysis, we explicitly account for cohort and business cycle effects that have taken place within the sample period (1978-2000).

Moreover, we allow for a reasonably flexible relationship between consumption or income on the one hand and age, year of birth and time on the other. We adopt the following empirical specification for consumption or income[3] (denoted by x_{it}):

$$\ln x_{it} = m_1\left(cohort_i\right) + m_2\left(age_{it}\right) + \phi_t D_t + \varepsilon_{it} \tag{4.1}$$

$m_1\left(cohort_i\right)$ is a linear spline with nodes at 1915, 1925, 1930, 1935, 1940, 1945, 1950, 1955, 1960, 1965, and 1970. $m_2\left(cohort_{it}\right)$ is a linear spline with nodes at 25, 30, 35, 40, 45, 50, 55, 60, 65, 70 and 75 years of age.[4] ε_{it} is a random error term. We

---

[2] This rental value has been assessed by real estate agents.

[3] The same empirical specification forms the basis of all figures we report in this research. Note that we did not constructed logs of the demographic variables and the participation rates.

[4] We have experimented with more flexible specifications (i.e., additional nodes in the spline

have computed standard errors that are robust to the presence of heteroscedasticity and of within cohort correlations. Note that our empirical specification precludes interactions between age and cohort effects.

Age, cohort and calender year effects are not separately identified because they are linearly dependent (i.e., age + cohort = calender year). We follow a method of disentangling age, period and cohort effects that has been proposed by Deaton and Paxson (1994). This method is used in many papers (e.g., Attanasio (1998) and Fernández-Villaverde and Krueger (2007)). The method boils down to imposing additional assumptions on the time dummies. The time dummies are constructed to be uncorrelated with a time trend and normalized such that they add up to zero (Deaton and Paxson 1994). This assumption for example implies that all 'trending' we find is interpreted as a cohort effect.

Figure (4.1) shows the predicted logarithm of real income as a function of age and year of birth respectively. Figure (4.1) reveals that real household income is strongly hump-shaped over the life cycle. Between age 21 and 25 the yearly growth rate in household income is equal to a rather extreme 12.5 % and between age 25 and 30 to 4.1 %. After, the annual growth rate is lower and household income tops at age 45. We more or less follow Fernández-Villaverde and Krueger (2007) by defining the size of the hump as the difference between income at age 21 and income at the top (age 45). As can be inferred from figure (4.1), the hump in log household income is about 0.92. This means that between age 21 and 45 income increases by 150 % (=$(e^{0.92} - 1) * 100$).[5] Between age 50 and 80 income steadily decreases with 50 % (=$e^{-0.69} - 1) * 100$). We would also like to stress that our estimate of the life cycle profile of income is rather precise (i.e the standard errors of our estimated age coefficients are small). In other words, our finding of a hump-shaped age-income profile is not merely due to sampling error.

We do not find important cohort effects for income, implying that real income *per household* has not increased on average over the sample period. One would expect that younger generations have a higher income than the older ones (ceteris paribus) (see e.g. Kapteyn et al. (2005)). However, it should be realized that households have decreased in size over the sample period. Where real income *per household* has not increased over generations, real income *per capita* did. We return to this issue in the next sections. The time pattern of the estimated year dummy coefficients (not reported here) basically follows the business cycle.[6]

The second set of graphs we show in figure (4.2) are the age and cohort profiles

---

function of age). This however, barely affected our empirical results.

[5] In figure (4.1), one should not interpret the reported values of ln(income) such 3.56 at age 45. The figure is only informative about the shape of the age income profile. The same caveat can be made for all other figures presented in this chapter.

[6] The complete set of estimation results are available from the authors upon request.

Figure 4.1: *Log household income as a function of (a) age (left panel) and (b) the year of birth of the household head (right panel).*

of both non-durable and durable consumption per household. Both age-consumption profiles show an important hump shape. The hump in ln(non-durable consumption) is approximately 0.84 and tops –like income– around age 45. Notice that this hump is slightly smaller than for household income. Interestingly, after age 50 non-durable consumption decreases at a slower pace than income. As we said before, our nondurable consumption measure includes rents and the imputed rental value of an owned dwelling. If we exclude these two items, we obtain an age profile which is very similar to that of income especially after age 50 and to the one obtained by Fernández-Villaverde and Krueger (2007) (see figure (2) of their paper). Figure (4.2) also shows that, ceteris paribus, younger cohorts consume more nondurable goods than the older ones. Non-durable consumption has increased from the 1905 cohort until the 1930 and stays approximately flat after.

The hump in durable consumption is about 0.5 but, –in contrast to income and non-durable consumption– drops down much further after age 50.[7] We find a surprising decrease in durable expenditures per household over generations. We find that in real terms, and when holding all else equal, younger cohorts spend less on durable items than older cohorts.

The age-consumption profiles are hard to reconcile with basic specifications of the life cycle model. If households are aware of the hump shape in income, households would typically borrow for consumption when income is low and save when income is high. This in order to smooth the (expected) marginal utility of consumption over the life cycle. The fact that consumption seemingly *tracks* income seems inconsistent with this argumentation. The evidence from figure (4.1) and (4.2) therefore sugges-

---

[7] Fernández-Villaverde and Krueger (2007) obtained a similar result.

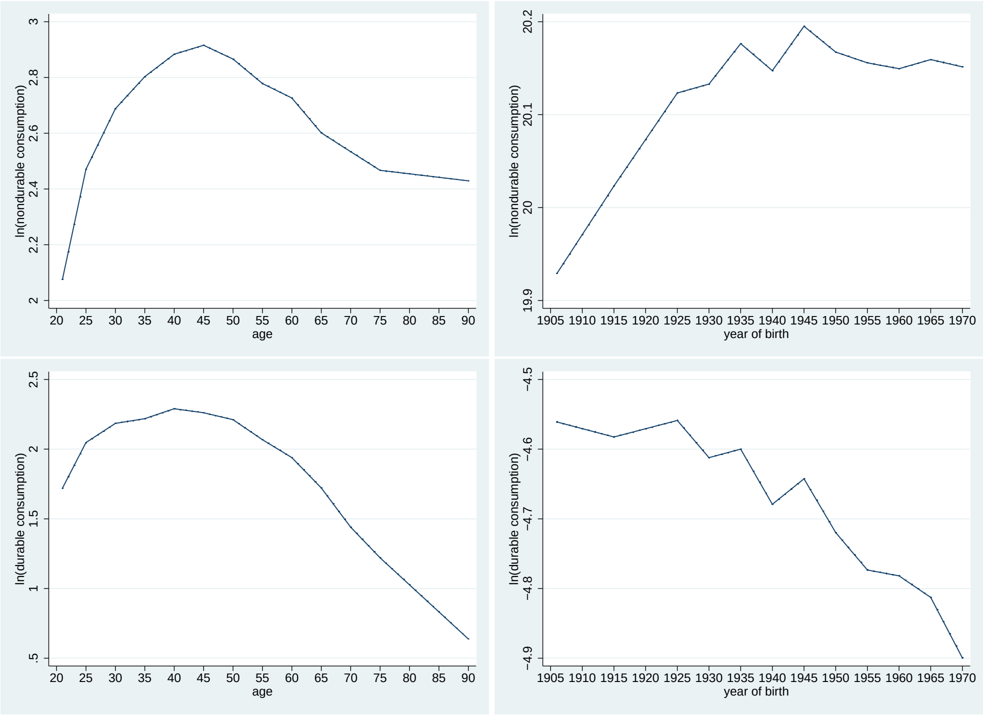


Figure 4.2: *Log consumption as a function of a) age and b) year of birth of the household head.*

tively rejects the basic specifications of the life cycle model. The hump shape in consumption however, is not necessarily inconsistent with more elaborate versions of the life cycle model.

4.3.1 Controlling for household composition

There are many competing explanations for the hump shape in consumption. Life cycle theory predicts that the expected marginal rate of substitution between consumption now and in the future should be equal to a constant (that is a function of the discount rate and the interest rate). If –for example– non-separabilities between consumption, leisure and household composition are important however, the marginal rate of substitution between consumption now and in the future is a function of household composition and/or labor supply variables. It has been frequently argued that the hump in consumption should, at least for a part, be attributed to the hump in household size [e.g., Attanasio and Weber (1995)].

In the Netherlands –and also elsewhere in the developed world– household size has decreased towards the end of the previous century. In part this is due to an increase in the number of single person households. In addition, we see that couples have less children now, than they had half a century ago. Figure (4.3) plots the number of adults within the household, as a function of age and year of birth respectively. The age profile is fairly flat until age 50 after which the probability of one of the household members dying becomes increasingly important. The decreasing cohort effects on the other hand, reflect the steady increase in the number of single headed households. The number of adults within a household has decreased by about 0.6 over generations.

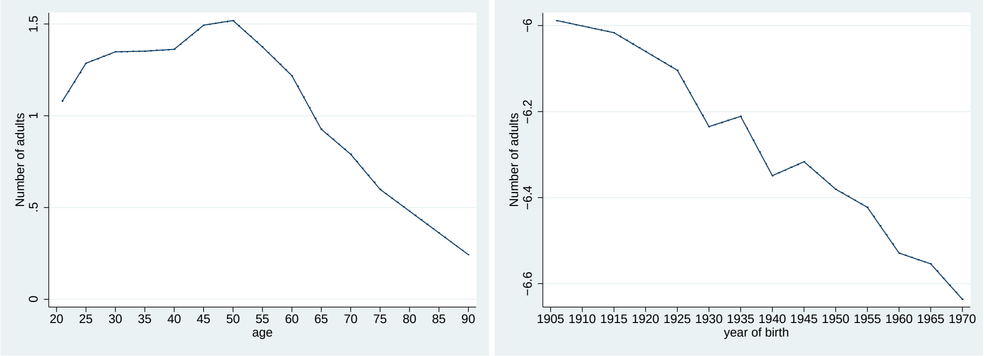


Figure 4.3: *Number of adults as function of age year of birth of the household head.*

Figure (4.4) documents the well-known hump shape in the number of children that peaks around age 35 to 40. After this age children start moving out of the household of their parents and start their own. The cohort pictures reveal another important –some would say alarming– feature of Dutch household data. Starting of with the 1930 cohort, younger cohorts start getting less and less children. We document a decrease of more than one child per household.

Figure (4.3) and (4.4) give helpful insights for interpreting the the consumption profiles of figure (4.2). For a given level of consumption increasing household size increases the marginal utility of consumption. Equating expected marginal utilities over time therefore, does not necessarily imply equating consumption levels over time.

We follow Fernández-Villaverde and Krueger (2007) and allow for household size by constructing equivalent consumption levels. Equivalent consumption is constructed by dividing household consumption by the modified OECD equivalence

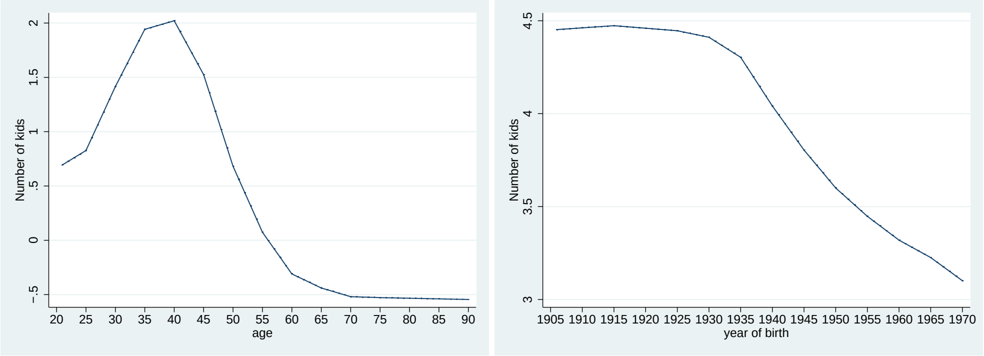


Figure 4.4: *Number of children as function of age and year of birth of the household head.*

scales as proposed by De Vos and Zaidi (1997).[8] The modified OECD scale is a household specific index that assigns a 1 to the first household member, 0.5 to every additional adult and 0.3 to each child. A couple with two children therefore gets assigned a equivalence scale of $1 + 0.5 + 0.3 + 0.3 = 2.1$]. The idea is that households maximize expected life time utility by allocating *equivalent* consumption efficiently.

The age and cohort profiles of equivalent income, equivalent non-durable and equivalent durable consumption are shown in figure (4.5). From comparing this figure with figures (4.1) and (4.2) we immediately see that allowing for household size (in the way we do it here) matters quite a bit for the age and cohort profiles. The age profiles are much flatter and no longer peak around age 45, but much later at around 55 or 60, typically after the children have moved out of the household. Between age 60 and 65 we see that income per equivalent adult decreases with only 1.6 percent per year.[9] Early retirement has also a small effect on equivalized income. After age 65 drops a bit further but the rate of decline is still rather small.

The drop in income around retirement does not translate into a drop in non-durable consumption.[10] Hence, after correcting for household composition, non-durable consumption does not drop after retirement. These findings are quite different from Fernández-Villaverde and Krueger (2007). Equivalizing income and non-

---

[8] After a careful comparison of different scales, Fernández-Villaverde and Krueger (2007) have decided to use an equivalence scale which is similar to ours.

[9] Kalmijn and Alessie (2008) found a relationship between equivalized income and age which is very similar to the one presented in figure (4.5). This is a comforting result because Kalmijn et. al. use tax record data in their analysis. Contrary to survey data measurement error in administrative data is not so much an issue.

[10] This is not true for a nondurable consumption measure which excludes housing. Even the drop in this consumption measure is rather small.

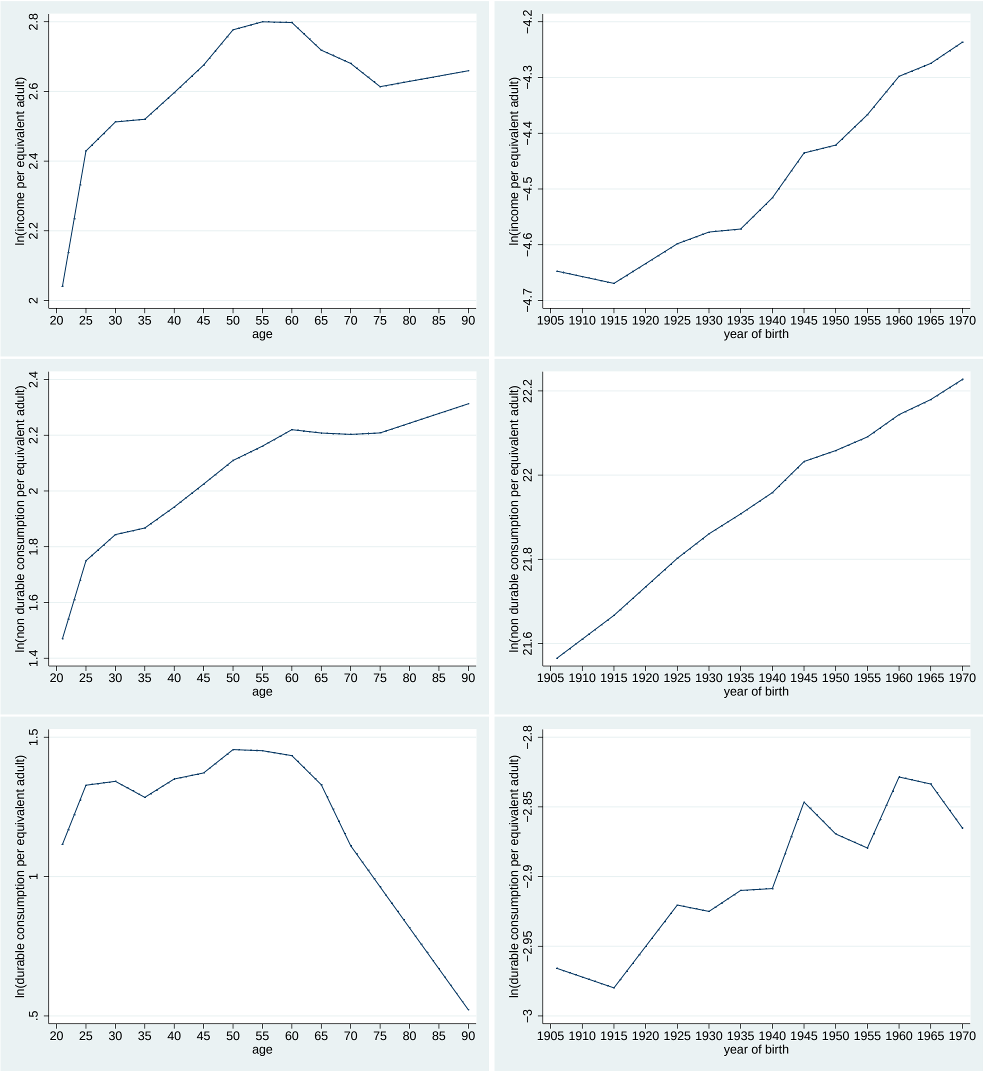


Figure 4.5: *Predicted* equivalent *income, non-durable consumption, and durable consumption as function of age and year of birth of the household head.*

durable consumption had, in their case, not a dramatic impact on the shape of the age profiles. According to their results, nondurable consumption per equivalent adult drops considerably after age 50 (see figure (5) of their paper). As compared with the US, the Netherlands has a rather generous pension system. Such institutional differences between the US and the Netherlands might affect consumption behavior differently. Although, on the basis of this evidence alone, we cannot exclude the possibility that other issues are at play here.

Very early in the life cycle (between age 21 and 25), durable consumption expenditures per equivalent adult rises sharply. This can possibly be attributed to binding liquidity constraints. From age 30 until 60 however, nondurable consumptions remains almost constant (it rises with a mere 10 percent over a period of thirty years). So, changes in household size are able to explain the hump in durable expenditures *up to* age 60. Yet, the drop in durable consumption *after* age 60 is striking. For as long as there is a consumption-retirement puzzle it seems to be related too consumer durables. Note however, that we do not observe the service flow of the stock of consumer durables, but merely expenditures.

Allowing for household composition above, also identifies an –intuitive– increase in in the cohort profiles. Where we found negative cohort effects for durable expenditures per *household*, we find a positive cohort effect per *equivalent adult*. This is easily explained by the strong cohort effects in household size. Where, on average, household expenditures on durables decreased over generations, household size has decreased even quicker. The net effect is an increase of durable consumption per equivalent adult.[11]

4.4 Explaining the *rest* of the hump

Even after controlling for household demographics we still find that non-durable consumption seems to track income in the first phase of life, where durable expenditures stay constant. Then when income drops, durable expenditures drop as well. Yet, in this case non-durable expenditures remain unaffected. The next set of graphs show the participation rates (fulltime or part time) of the head of the household and the partner respectively.

Here also we see a clear hump shape in both figures. (In addition we observe a slight dip around age 30 for the participation rate of the partner, perhaps reflecting the temporary decrease in participation rate to take care of young children at home.) We do not find cohort effects with the head of the household. For the partner (mostly women) we estimate that the participation rate picked up by about 0.2 (or 20% points) from the 1925 cohort to the 1965 cohort.

Not surprisingly the participation rate (for the head of the household) drops at about the age when income drops, i.e., around age 60. This indicates that at least

---

[11] Note that we allow for household size in one specific way. If we allow for household size in a more flexible way we obtained similar results for non-durables and income (output available on request). A more flexible specification of household size eliminated the cohort effects for durables.

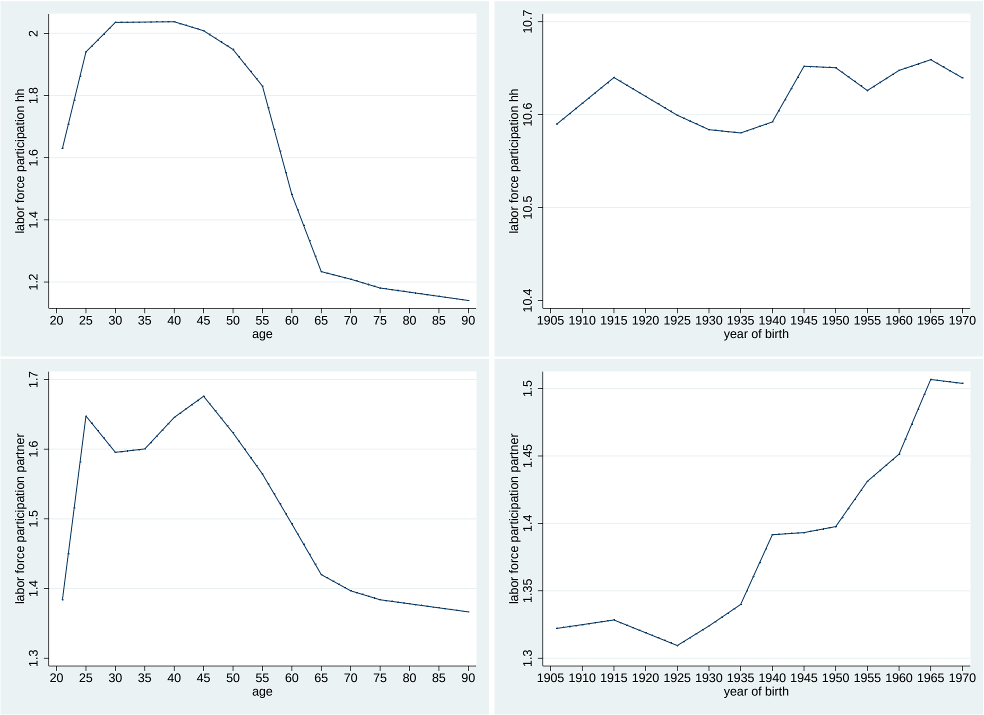


Figure 4.6: *Labor force participation rates as a function of (a) age (left panels) and (b) the year of birth (right panels). Household head (top panels), partner of household head (bottom panels).*

part of the decline in durable expenditures is work related. It makes sense that after leaving the labor force, cars –an important durable item– are replaced less frequently. However, an equally valid explanation could be possible nonseparabilities between consumption and leisure. If durable consumption and leisure are substitutes an increase in leisure will lead to a decrease in consumption.

It is less straightforward to rationalize the fairly steep increase in non-durable consumption per equivalent adult. The usual suspects put forward in the literature are borrowing constraints, or intertemporal nonseparabilities (e.g., habit formation). Households are willing to borrow against future wealth, but banks would not lend them the money. However, fiercely binding liquidity constraints seems an incomplete explanation for this empirical phenomenon. A more intuitive explanation is that people form habits over the course of life. Consumption as a result, is worth less in marginal utility terms around age 50 than around age 20. In order to overcome this depreciation of consumption households find it optimal to slowly increase consumption over the life cycle. A third candidate is that young cohorts are simply too uncertain about how their income is going to develop. A precautionary mechanism induces households to *rationally* give up on the option to borrow against expected future income increases.

4.5 Summary and conclusion

In this research we have used Dutch data to document the age profiles of consumption, with special emphasis on the distinction between expenditures on durables and nondurables. We find that the relationship between (non)durable expenditures and age is clearly hump shaped with a top at age 45. The age profile of nondurable consumption changes dramatically when we account for changes in household demographics. Although after such a correction nondurable consumption does not decline after retirement, nondurable consumption seems to track income at the early stages of life cycle. These findings do not seem to be in line with the theoretical predictions of a standard lifecycle model. In the previous section we have suggested some explanations for this finding but clearly more research is needed on this issue.

Like Fernández-Villaverde and Krueger (2007) we find that durable consumption expenditures (per equivalent adult) drops sharply after retirement. In this research we only provided some hints of how to explain this phenomenon. Over the last decade, several papers appeared interested in answering the question of why consumption expenditures drop at retirement. Those studies almost exclusively focus on nondurable consumption patterns. It seems worthwhile to extend this research by focussing on (the timing of) durable consumption (expenditures).

Finally it should be mentioned that our findings differ in some respects from those of Fernández-Villaverde and Krueger (2007). Using US data, they report as one of their main results the similar timing and size in the humps for expenditures on nondurables and durables, even after accounting for demographics. From figure (4.5) it is abundantly clear that this is not the case in the Dutch context. An interesting research question is to what extent differences in the institutional environment, such as pension and health insurance systems, between the two countries could explain those differences.

Chapter 5

Home production and the allocation of time and consumption over the life cycle

Abstract

This chapter estimates a model of female time allocation and non-durable consumption in an intertemporal utility maximization framework. We are using an extensive but relatively unexploited series of repeated cross sections from the Dutch B.O. consumer expenditure survey from Statistics Netherlands (1978-2000). Because male labor supply is known to respond rather inelastically to wage changes we condition on male labor supply in the analysis. We specify assumptions on domestic production technology that allow us to estimate labor supply elasticities that are consistent with non-separable preferences over consumption, leisure and a non-marketable domestically produced good, without needing time-use data. The empirical results can be summarized as follows: we estimate that female labor supply elasticities are about 1.7 if we take account of the intertemporal re-allocation of resources. This is about 50% larger than the elasticities we find in a static setting (they are about 1.1). Furthermore, we estimate intertemporal allocation parameters on a log linearized Euler equation using a synthetic panel with a large T dimension. These parameters are of reasonable size, but are imprecisely estimated. Moreover, we find that even after conditioning on demographics, income is a significant predictor for consumption growth in the period thereafter. This could be interpreted as evidence against the validity of our version of the life cycle model.

5.1 Introduction

Recent studies on time allocation emphasize the importance of domestic production (e.g., child care) within the context of female labor supply [see for example Becker (1965), Chiappori (1997), Apps and Rees (1997) and Apps and Rees (2003)].[1] These studies depart from the traditional dichotomous tradeoff between consumption and leisure and consider a tradeoff between three goods: consumption, leisure and a

---

[1] We use domestic production and home production interchangeably.

good that is produced within the home. In the Netherlands women who work, adjust the number of hours worked from about 30 to about 18 on average after having children. This clearly indicates the importance of this change of focus [see table (5.3). source: Statistics Netherlands expenditure survey].

Empirical studies dealing with consumption, leisure and home production in a simultaneous framework are rare. Data requirements for estimating such models are demanding as information on time-use must be combined with detailed information on wages, income and ideally consumption expenditures. Donni (2008) contributes to this problem by deriving properties of the domestic production function that subsequently allows for estimating and interpreting standard models of labor supply (i.e., where domestic production is mistakenly ignored).[2] Donni (2008)'s analysis hinges on the assumption that the domestically produced good is *marketable* (i.e., a close substitute can be obtained from the market).

It is likely that a mother's decision on hours of market labor is affected by the time she spends in child care. This would reject a key separability assumption that is often adopted in *baseline* empirical studies on labor supply (i.e., weak separability is rejected when the marginal rate of substitution between leisure and consumption depends on the level of home production). In this research we recognize and address this complication in a similar way as Donni (2008). We derive sufficient conditions on home production technology under which standard models of labor supply are well interpretable even if *non-marketable* home production is mistakenly ignored (i.e., preferences over home production are allowed to be nonseparable from leisure and consumption).[3]

Obviously without data on time-use it is difficult to make statements about how people substitute between leisure and domestic labor such that –in that sense– the model is not fully identified. Yet, the necessary parameters for predicting hours of market labor supply and for predicting labor supply elasticities can be estimated without explicit data on time-use. Under our assumptions on domestic technology standard preference parameters related to consumption and leisure should be reinterpreted as functions of preferences for consumption, leisure, the domestically produced good and domestic technology parameters. This outcome offers a theoretical rationale for interacting parameters of standard labor supply models with demographic characteristics such as the number of children in the household. Preferences for e.g.,

---

[2] Donni (2008) concludes: "a simple model of market labor supplies, which does not allow for domestic production, may conveniently fit the data if and only if (i) the profit function is additive or (ii) Engel curves are linear and the profit function has a particular, not necessarily additive form." "Quite importantly, however, the results ... crucially depend on the assumption that domestic goods are marketable (Donni 2008)".

[3] Providing necessary conditions –like Donni (2008) does for marketable home production– is beyond the scope of this study, but offers interesting and promising possibilities for further research.

child care are obviously dependent on the number of children you have.

We are furthermore considering the life cycle aspects of female labor supply behavior.[4] In this study we are endogenizing the intertemporal allocation of full expenditures within the context of an expected utility maximizing model. Modeling consumption and time allocation in a intertemporal framework seems important. We motivate this with a short example: *Young couples who plan to have children within a few years, foresee significant increases in child related expenditures. Moreover, when children are born, 'time' for leisure activities is scarce as children need to be looked after. To cover the expected increase in expenditures efficiently (i.e., by maximizing expected utility over the life cycle) households choose to increase current hours of market labor by sacrificing current leisure and increase savings. These households accumulate wealth that is used for consumption when children are born.* Static models typically do not capture this type of decision making.

This chapter presents two empirical novelties on Dutch micro data. We estimate a within period model on the allocation of female time and non-durable consumption that is consistent with intertemporal utility maximizing under uncertainty. This approach depends on the availability of combined information of labor supply data *and* data on consumption expenditures. Second, we estimate an Euler equation of non-durable consumption on a large T synthetic panel data set. Our empirical strategy is largely based on earlier work by Blundell and Walker (1986) as well as Blundell et al. (1994). Blundell and Walker (1986) estimate a model of household labor supply and consumption that is consistent with intertemporal two-stage budgeting under uncertainty. Blundell et al. (1994) study the life cycle allocation of household expenditures and the demand on different goods conditional on within period expenditure.

Blundell and Walker (1986) for example estimate a within period model that is consistent with life cycle theory. They do not however, estimate intertemporal preference parameters. Instead, they impute intertemporal preferences parameters for calculating life cycle consistent (or Frisch-) elasticities. In this chapter we estimate these parameters.[5] Furthermore, this study is the first that we know of that estimates an Euler equation of non-durable consumption on the basis of a Dutch large T micro data set.

---

[4] In this research we are interested in female labor supply behavior. Male labor supply, at least in the Netherlands, is known to respond relatively inelastically to wage changes (Theeuwes and Woittiez 1992, Evers et al. 2008). This may be due to important restrictions on the labor market. We choose not to worry too much about the complex institutional context of the decision process on male labor supply and use it as a conditioning variable in our analysis (Browning and Meghir 1991).

[5] In this chapter we often refer to intertemporal preference parameters as opposed to more specific concepts such as the elasticity of intertemporal substitution (a concept that is often directly estimated from Euler equations under the assumption of CRRA preferences). Due to the relative complexity of our model specification our intertemporal preference parameters do not exhibit such straightforward interpretations.

The Euler equation can be used to estimate preference parameters, but may also be used, simultaneously, to test the validity of the life cycle model. Ever since Hall (1978), many authors have tested the empirical tenability of the Euler equation of consumption, and also many rejections have been documented [see Browning and Lusardi (1996) for an overview]. The typical test studies the correlation between changes in consumption and initial income, or anticipated changes in income. Rejections of so-called *excess-sensitivity tests* have fueled the criticism on the credibility of the life cycle model. Others however, have shown that after controlling for demographics, the excess sensitivity disappears (Blundell et al. 1994). In this research we also perform such a test. We find that (log) changes in consumption are excessively sensitive to lagged income, even after controlling for demographic and labor supply variables. This result corroborates the findings of chapter 4 [also, Alessie and De Ree (2009)]. Chapter 4 finds that consumption is hump shaped over the life cycle, even after controlling for demographics. This empirical fact is hard to reconcile with the theoretical predictions of the Euler equation. Note that both chapter 4 and this chapter use the same data set.

Our main findings are summarized as follows: under specific, but reasonable assumptions on home production technology we are able to estimate a intertemporal model of female labor supply and consumption, without time-use data and without forcing undesirable separability assumptions on preferences. We have been estimating a within period model of female labor supply and non-durable consumption, conditional on male labor supply and demographic variables. We find no evidence for non-separable preferences over non-durable consumption and female non-market time.[6] Furthermore, we have estimated the Euler equation of non-durable consumption. After conditioning on demographics and male labor supply variables we still find excess sensitivity of consumption to lagged income. Finally, the intertemporal allocation parameters are are imprecisely estimated, yet of reasonable magnitude. Using the estimated parameters we have calculated female labor supply elasticities. We estimate a female labor supply elasticity of 1.1 in a static context. If intertemporal substitution of resources is taken into account, the elasticity is estimated to be larger, around 1.7. This result can be intuitively explained as follows: in a static context, an increase in current wages increases the price of current leisure compared to that of current consumption. This induces women to work more, in order to consume more (i.e., the static response). An increase in current wages however, also increases the price of current leisure compared to that of *future* leisure. This consequently induces women to increase working hours even more (i.e., the dynamic response).

The chapter is organized as follows: section (5.2) defines a theoretical framework

---

[6] Female non-market time is defined as female time other than market labor supply. Non-market time is the sum of hours of leisure and hours of domestic labor supply.

and derives the structural relationships that are used in estimation. Section (5.3) elaborates on the data set we use in estimation. Section (5.4) deals with the econometric techniques we need to estimate the parameters of interest and presents the regression results and the estimated elasticities. We report average within group elasticities. Groups were selected on the basis of age and on the presence of children within the household (Blundell and Walker 1986).

5.2 Theory

In the introduction we emphasize the need for a multi-period model to describe preferences. We assume that households maximize expected (exponentially discounted) utility over the life cycle, subject to a set of constraints. We follow Blundell et al. (1994) by defining instantaneous utility.

$$U_t = F_t \left(u_t \left(c_t^{ND}, l_{m,t}, l_{f,t}, x_t, \tilde{z}_{1t} \right), z_{2t} \right) + H \left(z_{3t} \right) \tag{5.1}$$

Instantaneous utility is a function of non-durable consumption c_t^{ND}, male leisure $l_{m,t}$, female leisure $l_{f,t}$, a non-marketable domestically produced good x_t and three vectors of (possibly endogenous) conditioning variables $\tilde{z}_{1t}$, z_{2t} and z_{3t}. These vectors may contain male labor supply, durable consumption, demographic composition etc., and may contain overlapping elements. The $\tilde{z}_{1t}$ variables are non-separable from the goods of interest and directly affect within period demand. The z_{2t} variables are weakly separable from goods of interest, such that they do not affect within period demand, but affect the intertemporal allocation of full expenditure. Variables in z_{3t} are additively separable from the goods of interest and therefore do not affect any (important) decision making directly. $H \left(z_{3t} \right)$ is therefore excluded from the analysis.

The value function $V_t \left(A_t \right)$ represents the value of all future expected discounted utilities as a function of wealth [like Zeldes (1989)].

$$\begin{aligned} V_t \left(A_t \right) = & \\ & \max_{c_t^{ND}, d_{m,t}, d_{f,t}, h_{f,t}} \Big[F_t \left(u_t \left(c_t^{ND}, l_{m,t}, l_{f,t}, x_t, \tilde{z}_{1t} \right), z_{2t} \right) + E_t \frac{1}{1+\delta} V_{t+1} \left(A_{t+1} \right) - \\ & \qquad \mu_{1ft} \left(-h_{ft} \right) - \mu_{2mt} \left(-d_{mt} \right) - \mu_{2ft} \left(-d_{ft} \right) \Big] \end{aligned} \tag{5.2a}$$

where

$$
\begin{aligned}
A_{t+1} &= (1+r_{t+1})\left(A_t + \tilde{I}_t^{other} + (T_f - l_{f,t} - d_{f,t})\, w_{f,t} + \right. \\
&\quad \left. (T_m - l_{m,t} - d_{m,t})\, w_{m,t} - p_t^{ND} c_t^{ND} - E_t^D\right) && (5.2b) \\
x_t &= x\left(d_{m,t}, d_{f,t}; \Theta\right) && (5.2c) \\
A_L &= 0 && (5.2d) \\
l_{m,t} &= T_m - \overline{h}_{m,t} - d_{m,t} && (5.2e) \\
l_{f,t} &= T_f - h_{f,t} - d_{f,t} && (5.2f)
\end{aligned}
$$

We model the period t decision making process of non-durable consumption c_t^{ND}, female hours of leisure $l_{f,t}$, male domestic labor supply $d_{m,t}$ and female domestic labor supply $d_{f,t}$. We condition on male hours of market labor supply $\overline{h}_{m,t}$ by including it in $\tilde{z}_{1t}$. Male and female domestic labor supplies are the only production factors in producing x_t and Θ is a vector of production technology parameters [see e.g., Apps and Rees (2003) who make the same assumption]. $d_{m,t}$ and $d_{f,t}$ are fully determined conditional on domestic technology, wages and the amount x_t that the household decides upon. Both hours of market labor and hours of domestic labor cannot be smaller than zero. μ_{1ft}, μ_{2mt} and μ_{2ft} are the respective Kuhn-Tucker multipliers associated with the nonnegativity constraints on $d_{f,t}$, $d_{m,t}$ and $h_{f,t}$. The non-negativity constraint on hours of male market labor supply is automatically satisfied as male labor supply $\overline{h}_{m,t}$ is a conditioning variable in the analysis. The expected future value of wealth $V_{t+1}\left(A_{t+1}\right)$ is discounted at rate δ.

The right hand side of equation (5.2b) are end of period t savings[7] that earn an interest r_{t+1} at the beginning of period $t+1$. The end of period t savings is the sum of period t assets A_t, other income $\tilde{I}_t^{other}$, male market labor income $(T_m - l_{m,t} - d_{m,t})\, w_{m,t}$, female market labor income $(T_f - l_{f,t} - d_{f,t})\, w_{f,t}$, minus non-durable expenditures $p_t^{ND} c_t^{ND}$ (where p_t^{ND} is the price of non-durable goods), minus durable expenditures E_t^D. We assume that there is no bequest motive (equation (5.2d)). The model differs from a standard life cycle model of labor and consumption by the incorporation of home production, where hours of leisure depend directly on the level of home production.

It is convenient to rewrite the above defined consumer problem by constructing a cost function for the domestically produced good x_t and subsequently substitute this into the original problem. The cost function represents the minimum cost of attaining x_t and is a function of market wages, domestic technology parameters and x_t itself. Under certain assumptions on home production technology (to be

---

[7] Until now we have used ‘savings’ as short for the intertemporal allocation of full expenditures. Here we use the term ‘savings’ in the usual interpretation.

spelled out in the next section) this procedure identifies an *implied price* of the home produced good as the cost function maybe written as $p_t^{HP}\left(w_{m,t}, w_{f,t}, \Theta\right) x_t$ where $p_t^{HP}\left(w_{m,t}, w_{f,t}, \Theta\right)$ is the implied price of the home produced good and is a function of male and female wages and technology parameters Θ. This representation of the cost function allows us to deal –at least in theory– with non-marketable home production x_t in the same as way we deal with any other good. A problem with this relatively general representation of the cost function is that we somehow need to measure home production and time-use data to estimate the parameters of the home production technology function. The following section elaborates on some of the principles of home production. Under some –to our judgement reasonable– assumptions on home production technology, standard models of female labor supply remain well interpretable without using time-use data.

5.2.1 Home production

We start the analysis on home production with discussing home production technology. In line with standard models of the firm, Apps and Rees (2003) for example introduce a home production function that is strictly quasi concave. Quasi concavity implies complementarities of both production factors (e.g., male and female time). In standard production theory, when capital and labor are the two key factors of production, it is intuitive that such complementarities exist. For home production –when male and female time are the only factors of production– the story is different. It is not obvious why a mother would be more efficient in walking children to school while her spouse is mowing the lawn. A domestic production model that assumes perfect substitutability of both factors of production seems more appropriate for domestic jobs like walking children to school or mowing the lawn. Note that perfect substitutability does not imply that both household members are equally efficient in producing, but only that the marginal rate of substitution between the factors of production is constant.

We note however that perfect substitutability is not satisfactory for all domestic jobs. A child may benefit from being raised by both of its parents such that complementarities between male and female time should be introduced. Which technology function gives a better fit to reality is ultimately an empirical question. Interesting progress could be attained by testing one against the other econometrically with time-use data. In this research we assume that the factors of production are perfect substitutes. In addition we assume that the production function exhibits constant returns to scale. Walking the children to school n times requires n time inputs, no matter the size of n.

For many types of domestic jobs both assumptions seem reasonable. Under the assumptions mentioned above, the production function has the following form:

$$x_t = x\left(d_{m,t}, d_{f,t}\right) = k \times \left(\pi d_{m,t} + d_{f,t}\right) \tag{5.3}$$

k is a measure of productivity of the household and π is measure of male productivity relative to female productivity. When $\pi = 1$ males and females are equally productive in producing x_t. The cost function is derived by minimizing the cost of production for a given level of output and non-negative domestic labor supplies:

$$\begin{aligned} &\text{cost}\left(w_{m,t}, w_{f,t}, x_t\right) = \\ &\quad \min_{d_{m,t}, d_{f,t}} \left[w_{m,t} d_{m,t} + w_{f,t} d_{f,t} \middle| x_t = k \times \left(\pi d_{m,t} + d_{f,t}\right), -d_{mt} \leq 0, -d_{ft} \leq 0\right] \end{aligned} \tag{5.4}$$

The cost of one hour of domestic labor are hourly market wages (i.e., the forgone revenue of one hour of market labor supply). The cost function conditional on our production technology (5.3) becomes:

$$\text{cost}\left(w_{m,t}, w_{f,t}, x_t\right) = p_t^{HP} x_t = \begin{cases} \frac{w_{f,t}}{k} x_t & \text{if } w_{f,t} < \frac{w_{m,t}}{\pi} \\ \frac{w_{m,t}}{k\pi} x_t & \text{if } w_{f,t} > \frac{w_{m,t}}{\pi} \end{cases} \tag{5.5}$$

The cost function consists of two parts. First, when female wages relative to female productivity is smaller than the male counterpart it is optimal for women to take care of all home production x_t that the household decides upon. Similarly, when male wages relative to male productivity is smaller then the female counterpart it is optimal for men to take care of all home production. The first representation of the cost function is probably the most appropriate for the Dutch situation, as on average, women earn less then men and also, most often, take care of the children at home (at least during office hours). Note that x_t could in principle be a vector that contains different kinds of home produced goods, each produced with technology similar to equation (5.3). The analysis could be extended for x_t being a vector of multiple domestically produced goods where domestic jobs are allocated between husband and wife in terms of their relative efficiency in producing. If all goods are produced using the technology defined by equation (5.3), the subsequent analysis does not change fundamentally.

In the next sections we study the most obvious case where women are cost efficient in producing the home produced good, but the exact same reasoning will apply to the case when men are cost efficient (i.e., we assume that $w_{f,t} < \frac{w_{m,t}}{\pi}$). Conventional models of labor supply are consistently estimated in both cases, but the interpretation of the preference parameters differs between the two. Under the assumption that women are cost effective in home production the cost function of the home produced goods boils down to the following simple representation:

$$\text{cost}\left(w_{m,t}, w_{f,t}, x_t\right) = p_t^{HP} x_t = \frac{w_{f,t}}{k} x_t \tag{5.6}$$

Male time in producing the home produced good equals zero, such that the production function (equation (5.3)) simplifies as follows:

$$x_t = k d_{f,t} \tag{5.7}$$

The key result of our assumptions on home production technology is that the *implied price* of the home produced good is a linear function of either male or female wages. Perfect substitutability of production factors (male and female time) and constant returns to scale of home production technology is sufficient to satisfy this property. Under these assumptions, the price of the home produced good is a linear function of either male or female wages on the whole positive wage domain. This fact has interesting implications. For example, as prices of $l_{f,t}$ and x_t move in parallel, Hicks' composite commodity theorem applies.

Deaton and Muellbauer (1980b) elaborate on the working of Hicks' composite commodity theorem:"...this is the composite commodity theorem, which asserts that if a group of prices move in parallel, then the corresponding group of commodities can be treated as a single good." Define a within period cost function $c\left(w, p^{HP}, p, u\right)$ governing preferences over leisure l, the home produced good x and a consumption vector q with respective prices w, p^{HP} and p. As a result from our assumptions on home technology the implied price of the home produced good is proportional to (male or female) wages: $p^{HP} = \frac{w}{k}$. The cost function therefore becomes $c\left(w, \frac{w}{k}, p, u\right)$ which –since k is a fixed parameter– can be written as a function of w, p and u alone: $c^*\left(w, p, u\right)$. Differentiating this cost function with respect to w yield the compensated demand for the composite good $\frac{\partial c^*(w,p,u)}{\partial w} = l + \frac{x}{k}$.[8] So, $l + \frac{x}{k}$ is the quantity of the composite commodity that corresponds to price w. "Since the cost function provides a complete picture of preferences, this demonstration shows that ... new preferences can be defined over q and $l + \frac{x}{k}$ and that these preferences lead to the same choices as the original ones (Deaton and Muellbauer 1980b)."

The composite good corresponding to female wages is the sum of female leisure and home production scaled by a productivity factor $l_{f,t} + \frac{x_t}{k}$. Using equation (5.7) it

---

[8]
$$\frac{\partial c^*\left(w, p, u\right)}{\partial w} = \frac{\partial c\left(w, \frac{w}{k}, p, u\right)}{\partial w} = c_1\left(w, \frac{w}{k}, p, u\right) + c_2\left(w, \frac{w}{k}, p, u\right) \times \frac{1}{k} = l + \frac{x}{k} \tag{5.8}$$
c_1 and c_2 are partial derivatives with respect to the first and the second element respectively.

is straightforward to show that the composite good equals female time that is spent outside the labor market, hence the sum of hours of leisure and hours of domestic labor supply. We define this variable as female 'non-market time' or n_t:

$$n_{f,t} = l_{f,t} + \frac{x_t}{k} = l_{f,t} + d_{f,t} \tag{5.9}$$

Female market labor supply is the complement of female non-market time: $h_{f,t} = T_f - n_{f,t}$. The above derivations show that under our technology assumptions, standard (static as well as dynamic) models of labor supply can be consistently estimated and interpreted even if time-use data is not available. The traditional dichotomous tradeoff between leisure and consumption should be reinterpreted as a tradeoff between non-market time and consumption. It is obvious therefore that the interpretation of the preference parameters as well as the estimated elasticities should also be reinterpreted. Preferences for consumption and non-market time are composites of preferences for consumption, leisure, home produced goods, and domestic productivity parameters. The change of focus also indicates that discriminating households with and without children deserves additional attention. Clearly, when couples become parents preferences for non-market time will change drastically. Preferences for the home produced good, female leisure and the productivity parameters cannot be identified separately without the appropriate data (i.e., time-use data, definitions of home produced goods).

5.2.2 A model of female time allocation

We simplify the original consumer problem described by equation (5.2) by incorporating the results on home production derived in the previous section. The new model deals with the life cycle allocation of non-market time and consumption.

$$\begin{aligned} V_t\left(A_t\right) &= \max_{c_t^{ND}, n_{f,t}} \Big[F_t\left(u_t\left(c_t^{ND}, n_{f,t}, z_{1t}\right), z_{2t}\right) + \\ &\qquad E_t \frac{1}{1+\delta} V_{t+1}\left(A_{t+1}\right) - \mu_{1ft}\left(-\left(T_f - n_{f,t}\right)\right) \Big] \end{aligned} \tag{5.10a}$$

subject to:

$$A_{t+1} = \left(1 + r_{t+1}\right)\left(A_t + I_t^{other} + \left(T_f - n_{f,t}\right) w_{f,t} - p_t^{ND} c_t^{ND} - E_t^D\right) \tag{5.10b}$$

$$A_L = 0 \tag{5.10c}$$

where I_t^{other} includes male income, hence $I_t^{other} = \tilde{I}_t^{other} + \overline{h}_{m,t} w_{m,t}$. z_{1t} includes male leisure. The non-negativity constraints on female domestic labor supply are satisfied by equation (5.4).

λ_t is defined as the marginal utility of wealth $V_t'(A_t)$. The conditions for a maximum of the above decision problem are:

$$1 \quad : \quad \frac{\partial F_t}{\partial u_t} \frac{\partial u_t}{\partial c_t^{ND}} - \lambda_t p_t^{ND} = 0 \tag{5.11a}$$

$$2 \quad : \quad \frac{\partial F_t}{\partial u_t} \frac{\partial u_t}{\partial n_{f,t}} - \lambda_t w_{f,t} - \mu_{1ft} = 0 \tag{5.11b}$$

$$3 \quad : \quad -(T_f - n_{ft}) \leq 0 \quad \text{with equality if } \mu_{1ft} > 0 \tag{5.11c}$$

$$4 \quad : \quad \lambda_t = E_t \frac{1 + r_{t+1}}{1 + \delta} \lambda_{t+1} \tag{5.11d}$$

μ_{1ft} is the Kuhn-Tucker multiplier associated with the inequality constraints on hours of female market labor supply. For the moment we assume that the Kuhn-Tucker multiplier is zero such that the nonnegativity constraint on hours of market labor supply is not binding. We account for this data selection in the empirical analysis by adjusting the likelihood function [see Blundell and Walker (1986)]. Equation (5.11d) presents the well-known result of intertemporal utility maximization models. In expectation, discounted marginal utility of wealth should be constant over time. This type of relationships are known as Euler equations and are used in section (5.4) to identify intertemporal allocation parameters.

Two-stage budgeting and the specification of preferences

The availability of panel data (or synthetic panel data) is not a prerequisite for estimating the parameters of a model that is *consistent* with intertemporal maximization under uncertainty. By combining condition (5.11a) and condition (5.11b) with the within period budget constraint one is able to construct a model of labor supply that is consistent with expected utility maximization under uncertainty [e.g., Blundell and Walker (1986)]. We also use this property in this study. This allows us to estimate all the relevant parameters of the model in two consecutive steps.

This two-stage identification strategy relates to the idea of two-stage budgeting where in the first stage, households allocate life-time resources over consecutive periods, fixing full expenditure in every period. In the second stage, households allocate full period t expenditures over within period goods. Full expenditures in our case is the total expenditures on the goods of interest, female non-market time and non-durable consumption. The attractiveness of the two-stage budgeting idea is that the parameters of the instantaneous utility function $u(\cdot)$ are identified on the basis of within period prices, within period demands, and within period full expenditures only. As a consequence, the parameters of $u(\cdot)$ can be estimated using only cross sectional data. In order for this to work, however, preferences need to be (weakly) separable over time.[9]

The second stage model is constructed by combining the first order conditions (5.11a) and (5.11b) with full expenditures on the goods of interest y_t. y_t operates as a suitable conditioning variable capturing "future anticipation and past decisions (Blundell and Walker 1986)". The parameters of F are not identified in the second stage as the terms that include F cancel out by combining (5.11a) and (5.11b).

In this research we derive the second stage demand system from an indirect utility function. The *conditional indirect utility function* is defined as follows:

$$\begin{aligned} & F\left(\Psi_t\left(p_t^{ND}, w_{f,t}, z_{1t}, y_t\right), z_{2t}\right) = \\ & \qquad \max_{c_t^{ND}, n_{f,t}} \left[F\left(u\left(c_t^{ND}, n_{f,t}, z_{1t}\right), z_{2t}\right) \middle| y_t = p_t^{ND} c_t^{ND} + w_{f,t} n_{f,t}\right] \end{aligned} \tag{5.12}$$

It is not uncommon to define indirect utility functions (or cost functions) as opposed to direct utility functions to derive demand functions. Indirect utility functions exhibit the important advantage that they may be specified sufficiently flexible to accommodate important features of demand while at the same time, algebraic solutions for the structural relationships (e.g., demand equations) still exist. Direct utility functions do not exhibit both features simultaneously. However, to facilitate the estimation of the intertemporal preference parameters using Euler equations of non-durable consumption it can be profitable to work with direct utility functions [see e.g., Bean (1986) and Ziliak and Kniesner (2005)]. Such an approach sacrifices the possibility of deriving analytical solutions for the demand equations. Ziliak and Kniesner (2005) solve this inconvenience by estimating the parameters of a direct utility function (within the context of a life cycle model of labor supply) on the basis of within period marginal conditions, rather than demand systems.[10]

---

[9] In the empirical section (5.4) we find that consumption growth is excessively sensitive to lagged income measures. Excess sensitivity may have many causes of which intertemporal non-separability is one.

[10] In the empirical section of this chapter we were unable to reject the linear expenditure system in favor of more general specifications. For the linear expenditure system there exists a direct utility function. In section (5.4.2) we use this direct utility function to derive a Euler equation of non-durable consumption that is subsequently used for estimating the parameters of F.

We are specifying the following functional form to operationalize equation (5.12):

$$F\left(\Psi_t\left(p_t^{ND}, w_{f,t}, z_{,1t}, y_t\right), z_{2t}\right) = \frac{1}{1-\rho}\left(\frac{y_t - a\left(p_t^{ND}, w_{f,t}, z_{1t}\right)}{b\left(p_t^{ND}, w_{f,t}, z_{1t}\right)}\right)^{1-\rho} \times \exp\left[\alpha' z_{2t}\right] \tag{5.13}$$

The functional form to organize intertemporal preferences $F(\cdot)$, is the C.R.R.A. (constant relative risk aversion) utility function. This specification is popular, because it allows for some important aspects of (intertemporal) behavior such as, for example, the precautionary motive [see (Browning and Lusardi 1996) for an overview]. The parameter ρ determines the curvature of the utility function and is a measure of risk aversion. When ρ is large, households display high aversion to future utility losses and indifference to future utility gains.

$a(\cdot)$ and $b(\cdot)$ are homogenous of degree 1 in within period prices and α is a vector of parameters:

$$\begin{aligned} a\left(p_t^{ND}, w_{f,t}, z_{1t}\right) &= \gamma_c(z_{1t})\, p_t^{ND} + \gamma_n(z_{1t})\, w_{f,t} + 2\gamma_{nc}\sqrt{p_t^{ND} w_{f,t}} && (5.14)\\ \ln b\left(p_t^{ND}, w_{f,t}, z_{1t}\right) &= \beta_c(z_{1t}) \ln p_t^{ND} + \beta_n(z_{1t}) \ln w_{f,t} && (5.15) \end{aligned}$$

The model nests the linear expenditure system if $\gamma_{fc} = 0$. The variables z_{1t} are allowed to affect within period decisionmaking directly. We allow for this dependence in roughly the same way as in Blundell and Walker (1986). The cost of living parameters (γ_n and γ_c) are interacted with family size (i.e., the number of household members fs). During demographic transitions preferences for the home produced good will change such that interacting the preference parameters with demographic variables is important.

$$\begin{aligned} \gamma_n &= \gamma_n^0 + \gamma_n^{fs} \cdot fs && (5.16)\\ \gamma_c &= \gamma_c^0 + \gamma_c^{fs} \cdot fs && (5.17)\\ \gamma_{nc} &= \gamma_{nc} && (5.18) \end{aligned}$$

γ_{nc} is not parameterized. The β parameters are interacted with age (age), male hours of market labor supply ($\overline{h}_m$) and three dummy variables $D1$, $D2$ and $D3$. The three dummy variables are one when the youngest child is below 6, between 6 and 12 and between 12 and 18 years of age respectively. The presence of young children is expected to strongly affect preferences for *non-market time*. We obtain:

$$\begin{aligned}\beta_n &= \beta_n^0 + \beta_n^{D1} \cdot D1 + \beta_n^{D2} \cdot D2 + \beta_n^{D3} \cdot D3 + \beta_n^{age} \cdot (age - 40) + \\ &\quad \beta_n^{h_m} \cdot \overline{h}_m \end{aligned} \tag{5.19}$$

$$\beta_c = 1 - \beta_n \tag{5.20}$$

The adding up constraint imposes $\beta_c = 1 - \beta_n$.

From applying Roy's identity to equation (5.13) one can derive the within period demand equations for nondurable consumption and female non-market time. One of the two demand equations can be dropped from the analysis without loss of generality (i.e., an implication of the adding-up constraint). In this research we focus on the demand equation for female non-market time, or actually, on its complement: female market labor supply. The female market labor supply function is subsequently constructed by subtracting demand for non-market time from total time endowment T_f:

$$h_{f,t}^{\text{uncomp.}} = T_f - n_{f,t} =$$

$$= -a^*_{w_{f,t}} - \frac{b_{w_{f,t}}}{b}\left[y_t^* - a^*\right] \tag{5.21}$$

$$= -\gamma_n^* - \gamma_{nc}\sqrt{\frac{p_t^{ND}}{w_{f,t}}} - \frac{\beta_n}{w_{f,t}}\left[y_t^* - \gamma_c p_t^{ND} - \gamma_n^* w_{f,t} - 2\gamma_{nc}\sqrt{p_t^{ND} w_{f,t}}\right] \tag{5.22}$$

where

$$a^*\left(p_t^{ND}, w_{f,t}, z_{1t}\right) = \gamma_c\left(z_{1t}\right) p_t^{ND} + \gamma_n^*\left(z_{1t}\right) w_{f,t} + 2\gamma_{nc}\sqrt{p_t^{ND} w_{f,t}} \tag{5.23}$$

$$\gamma_n^* = \gamma_n - T_f \tag{5.24}$$

$$y_t^* = y_t - T_f w_{f,t} = p_t^{ND} c_t^{ND} - h_{f,t} w_{f,t} \tag{5.25}$$

Due to our assumptions on preferences we do not need to specify total time endowment T_f as it is subsumed in one of the parameters. $a^*_{w_{f,t}}$ and $b_{w_{f,t}}$ are partial derivatives with respect to female wages of a^* and b respectively. y_t^* is our *new* measure of full expenditures. We estimate the parameters of Ψ using the relationship defined by (5.22) in section (5.4.1).

Compensated demands are obtained by inverting the conditional indirect utility function and by writing y_t^* as a function of prices and within period utility (i.e., the conditional cost function) and substituting this into equation (5.21):

$$h_{f,t}^{\text{comp.}} = -a_{w_{f,t}}^{*} - b_{w_{f,t}}\overline{U} \tag{5.26}$$

where $\overline{U}$ and is a measure of within period utility. Note that the parameters of F do not enter the compensated nor the uncompensated demand functions.[11]

The demand functions (5.21) and (5.26) are consistent with life cycle theory. The elasticities however, are conditional on full expenditures and within period utility respectively and therefore do not account for the intertemporal reallocation in response to a price (e.g., a wage) change. That is, households might not hold full expenditures constant after a change in wages, but in fact adjust full expenditures in response to a change in wages. A temporal increase in wages would increase labor supply conditional on y_t^*, but indeed, the temporal nature of the wage increase also makes *future* non-market time cheaper with respect to non-market time today. This would induces women to increase current labor supply even more. More specifically, our model assumes that y_t^* responds to within period price changes in such a way that marginal utility of wealth (λ_t) stays constant over time (in expected discounted terms). Demand functions that exhibit this property are called Frisch (or λ-constant) demand functions. It is intuitive that Frisch demand elasticities can be estimated only after having acquired information on the structure of intertemporal preferences F.

Life cycle consistent (or Frisch-) demand functions are obtained by writing marginal utility of wealth λ_t as a function of full expenditures and by substituting this relationship into the uncompensated demands:[12]

$$h_{f,t}^{\text{Frisch.}} \quad = \quad -a_{w_{f,t}}^{*} - b_{w_{f,t}}\left[\left(\frac{b\lambda_t}{\exp\left[\alpha' z_{2t}\right]}\right)^{-\frac{1}{\rho}}\right] \tag{5.27}$$

Price elasticities derived from equation (5.27) incorporate the effects of intertemporal reallocation of full expenditures on the goods of interest and lead to interesting deviations from standard uncompensated elasticities. Note however, that Frisch de-

---

[11] Within period utility U is defined as $F\left(\frac{y_t^*-a_*}{b}, z_{2t}\right)$. $\overline{U}$ is defined as $F^{-1}\left(U, z_{2t}\right)$.

[12] We can show that:

$$\begin{aligned}\lambda_t &= \frac{\partial}{\partial y_t^*}\frac{1}{1-\rho}\left(\frac{y_t^* - a^*\left(p_t^{ND}, w_{f,t}, z_{1t}\right)}{b\left(p_t^{ND}, w_{f,t}, z_{1t}\right)}\right)^{1-\rho} \times \exp\left[\alpha' z_{2t}\right] \\ &= \left(\frac{y_t^* - a^*\left(p_t^{ND}, w_{f,t}, z_{1t}\right)}{b\left(p_t^{ND}, w_{f,t}, z_{1t}\right)}\right)^{-\rho}\frac{1}{b\left(p_t^{ND}, w_{f,t}, z_{1t}\right)} \times \exp\left[\alpha' z_{2t}\right]\end{aligned}$$

mands are hard to interpret if, for example, liquidity constraints are binding, or, more generally, if the life cycle model is not a valid representation of preferences.

Blundell and Walker (1986) fix the parameters of F at a "reasonable" values (i.e., $\rho = 1$ such that $\frac{1}{1-\rho}(\cdot)^{1-\rho} = \log(\cdot)$). In this research we estimate ρ by estimating an Euler equation of non-durable consumption. We return to this issue in section (5.4).

5.3 Data

We estimate the parameters of our model using a time series of cross-sections drawn from the public-use files of the Statics Netherlands B.O. consumer expenditure survey. The survey has collected yearly data of around 2000 households from 1978 to 2000.[13] The study relates information on household's state, such as income, number of children, etc., to expenditure on a detailed set of consumption goods, services and taxes. Moreover, we have information on both male and female hours of paid labor for the 1988 to 1991 waves. All waves, from 1978 to 2000, contain information on whether both male and female work full-time, part-time or not at all.

For estimating the parameters of the within period model (the consumption - non-market time tradeoff conditional on within period full expenditures) we use the four waves that contain complete information on hours of labor and consumption (i.e., the 1988 to 1991 waves). We break down consumption expenditures into two broad categories: durable- and non-durable consumption goods. Non-durable consumption goods fully depreciate within one period. Durable consumption goods depreciate at a slower rate. Non-durable consumption goods are food, clothing, rent (as well as imputed rents for house owners). Durable consumption expenditures are investments on education, cars, furniture, refrigerators, etc. Durable expenditures is excluded from the analysis such that we effectively assume that preferences for durable consumption are separable from the key variables under study.

For estimating the parameters of F we estimate a dynamic model on a synthetic panel data set that is constructed out of all available waves of data set. It has been shown that dynamic models can be estimated using series of repeated cross-sections (e.g., Moffitt (1993) and Verbeek and Vella (2004)). Rather than using individual/household data, we construct cohort averages of the relevant variables and use these in estimation. Using synthetic panels however introduces sampling error, because cohort averages are just *estimates* of cohort means. To minimize the sampling error cohort averages should be constructed out of a sufficiently large group of individual households. Therefore, cohorts are defined on a five year interval (e.g.,

---

[13] From 1978 to 1987 the data set contained of around 2500 observations per year. From 1988 to 2000 of around 1700 observations per year, where 1991 is an negative outlier with 900 observations per year

households of which the household heads are born between 1950 and 1954, belong to the same cohort). Sampling weights are used to improve the representativeness of the cohort averages. We end up with 15 cohorts measured over 23 year (=345 observations). On average the number of households per cohort-time average is 185, with about 5% of the cohort/time averages that are constructed out of less than 50 individual households.

Three issues with missing data occurred. Most of our waves (i.e., all of them except for the 1988-1991 waves) lack information on hours of market labor supply of both adult members of the household. Fortunately, all waves contain information about the full-time or part-time labor status of both adult household members. These binary variables appear to be very good predictors for hours of labor. With the information on full-time or part-time labor status we have imputed hours where they were missing. To obtain estimates on hours we use a *hotdeck* imputation procedure.[14] Note that the imputed hours are merely used as a conditioning variable in estimating the Euler equation and not to estimate the within period labor supply model.

Second, the 1989 wave lacks information on tax expenditures other than income tax. Missing data is imputed by means of a regression analysis using data of the adjacent years, 1988 and 1990. i.i.d. error terms are added to the predicted values to preserve the variability in the data. Third, the data set does not contain explicit information on hourly wages. Evidently, wages are an important argument in a female labor supply model. We construct net wages by dividing female net labor income by the number of hours of labor. As hours of labor are likely to be measured with error, wages, consequently, are infected with the same measurement error. This is problematic not only because the measurement error leads to imprecisely estimated wages, but it also produces a spurious negative correlation between hourly wages and hours of labor. We solve the measurement error problem by instrumenting the (constructed) wages[15] and using the predicted wages in the labor supply regression.

5.3.1 Stylized facts at the macro (and at the micro) level

Before we estimate the parameters of the model we present some of the important stylized facts of the data. This improves intuition about the process that drives economic outcomes and checks for possible irregularities in the data. First we check whether the consumption series of the expenditure survey (our data set) corresponds to aggregate consumption series of the national accounts. Moreover, we check whether the predicted hours of labor exhibit the well-known upward trend that characterizes

---

[14] The hotdeck procedure takes random draws from variable x from appropriate sub samples of the data and allocates them randomly to households where variable x was missing.

[15] There is no reason to worry about endogeneity problems in this wage equation. The interest is not to identify causal relationships regarding wages, merely to get rid of the measurement error.

the development of female hours of labor supply over the sample period.

We have compared our data from the expenditure survey with other measures of consumption. To do so, we have constructed aggregate consumption series from the expenditure survey and plotted it together with macro economic figures on consumption from the national accounts in figure (5.1A.). For making both series comparable, data from the expenditure survey is corrected for population growth and change in family size.[16] The (corrected) measure of total consumption from the expenditure survey matches the aggregate measure of consumption from the national accounts very well such that there is no reason to reject one measure over the other in terms of reliability. When we impute hours of labor we assume that the expected hours of labor for the average full-time and part-time worker is constant over time. Hereby we attribute any visible change in hours of female labor activity over time to an increase in the number of females working full-time or part-time. The time series of figure (5.1B.) shows that an increasing number of women engage in market labor over time. Hence, the data confirms this well-known trend in the Netherlands.

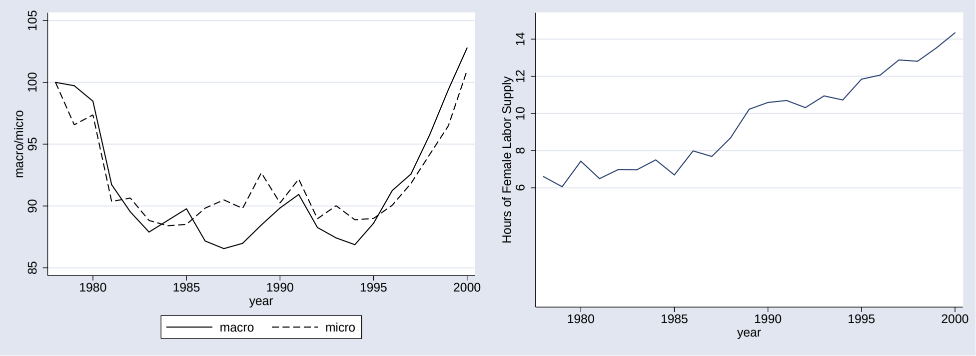


Figure 5.1: ***A.*** *Consumption National Accounts (macro) vs. Consumption Expenditure Survey (micro).* ***B.*** *Average Weekly Hours of Female Paid Labor, per household.*

On a micro scale we are interested in the relevancy of a life cycle model. We refer to the extensive description of the data in Alessie and De Ree (2009) and chapter 4 of this thesis. In chapter 4 we estimate life cycle profiles of durable and nondurable consumption, household income, and demographic and labor supply variables. Both types of consumption expenditures exhibit clear hump shapes over the life cycle (after controlling for birth cohort effects and time effects[17]). Furthermore, chapter 4 finds

---

[16] Family size decreases over the period under investigation. A constant level of consumption per household implies a rising level of consumption per capita.

[17] Note that age, cohort and time effects cannot be separately identified. In chapter 4 we follow Deaton and Paxson (1994) by imposing additional assumptions on the time dummies for identification.

that the humps can be only partly explained by a similar hump shape in household size. These results challenge the validity the "consumption smoothing" concept in a semi-parametric way. In addition, in this chapter we are also considering nonseparabilities with labor supply (or nonmarket time). In section (5.4.2) we propose a more formal test of our "extended" life cycle model.

Furthermore, we would like to refer to figure (4.4) of chapter 4. This figure shows the life cycle profiles and cohort profiles of the average number of children within households. It is apparent that younger cohorts have less children on average. It is likely that this phenomenon is related to changes in female labor supply, as younger cohorts also tend to work more (see figure (4.6) of chapter 4). Another feature of the data is that female hours of market labor supply show a dip around age 30 when, on average, children are born. This indicates the relevance of considering child care within the context of female labor supply and supports the focus of our model. We find no such pattern for male hours of labor, such that –when push comes to shove– women rather than men seem to take responsibility for child care.

5.4 The empirical analysis

Section (5.4) estimates all parameters of the life cycle model presented in section (5.2.2). In section (5.4.1) we are estimating the parameters of Ψ using the within period model of female labor supply. In section (5.4.2) we derive and estimate an Euler equation of non-durable consumption. Section (5.4.3) presents elasticities that are calculated on the basis of our regression results. We present compensated, uncompensated and Frisch elasticities of female labor supply and non-durable consumption, conditional on demographics and hours of male labor supply. Because the elasticities depend on demographics, we report average elasticities for groups with different demographic characteristics.

5.4.1 The second stage: the labor supply model

We estimate the labor supply model using the 1988 to 1991 waves (7663 observations). Information on time-use (other than market labor supply) is not available such that identifying the production parameter k or disentangling preference parameters for non-market time into preferences for leisure and the home produced good is not possible. Nevertheless, as we have shown in section (5.2.1), we are able to identify labor supply and consumption elasticities under certain assumptions on domestic technology. We estimate the parameters of the following regression model:

$$\begin{aligned} h_{f,it} &= -\gamma^*_{n,it} - \gamma_{nc}\sqrt{\frac{p^{ND}_{it}}{w_{f,it}}} - \\ & \frac{\beta_{n,it}}{w_{f,it}}\left[y^*_{it} - \gamma^*_{n,it}w_{f,it} - \gamma_{c,it}p^{ND}_{it} - 2\gamma_{nc}\sqrt{p^{ND}_{it}w_{f,it}}\right] + \eta_{it} \end{aligned} \quad (5.28)$$

The *it* subscripts indicate households i and time t [see e.g., equation (5.16), (5.17) and (5.19)]. η_{it} is an i.i.d. normal error term. For deriving this model we have imposed that the Kuhn-Tucker multipliers associated with the non-negativity constraint on female market labor supply μ_{1ft} are zero. As a consequence, this model does not apply to woman who do not work. We follow Blundell and Walker (1986) by selecting data on working females and by using a truncated regression method in estimation to adjust for this selection. We exclude non-working females, pensioners and the self-employed from the analysis, yielding 1906 observations with positive female market labor supply. The likelihood contribution of household i at time t becomes:

$$\phi_\eta\left(\eta_{it}\right)/P\left(h_{f,it} > 0\right) \quad (5.29)$$

$\phi_\eta\left(\cdot\right)$ is the normal density function of the disturbance term η_{it}. $P\left(h_{f,it} > 0\right)$ is the probability that a household is selected in the sample.

Real wages are constructed by dividing real after tax income by the reported hours (separate incomes for males and females are recorded). We use Stone price indices[18] to discount nominal income and non-durable consumption expenditures. Stone price indices are household specific price indices that correct for different rates of inflation for different bundles of goods [see for example Attanasio and Weber (1995)] [source: Statistics Netherlands. `http://statline.cbs.nl/`].

Male labor supply $\overline{h}_{m,t}$ and full expenditure y^*_t are simultaneously determined within the household and therefore endogenously related to female hours of labor supply. We carry out two Hausman type tests to test for weak exogeneity. The procedure simultaneously gets rid of the bias in the regression estimates (Smith and Blundell 1986). For this test we have performed two linear regressions where y^*_t and $\overline{h}_{m,t}$ are both regressed on the variables of the model and two separate vectors of

---

[18] Stone price indices are calculated as follows

$$\ln p^{Stone}_{it} = \sum_{j=1}^{K} w_{ji} \ln p_{jt} \quad (5.30)$$

where the price index of individual i depends on its income share spent on product category j, w_{ji} and the price index of that product category, p_{jt}.

Table 5.1: *Estimates of the within period model*

	(1)		(2)	
PARAMETERS	ESTIMATES	t-stat.	ESTIMATES	t-stat.
β_n^0	0.4513***	[20.9]	0.4525***	[21.0]
β_n^{D1}	0.0764***	[4.1]	0.0854***	[9.0]
β_n^{D2}	0.0435***	[3.6]	0.0483***	[5.6]
β_n^{D3}	0.0211***	[2.4]	0.0234***	[2.9]
β_n^{age}	0.0033***	[3.9]	0.0037***	[11.6]
$\beta_n^{h_{m,t}}$	-0.0009***	[-5.2]	-0.0010***	[-5.4]
γ_c^0	116.3694	[0.9]	102.4813	[0.9]
γ_c^{fs}	-126.7761***	[-2.8]	-119.3981***	[-2.8]
γ_n^0	-46.0052***	[-5.2]	-48.1035***	[-6.1]
γ_n^{fs}	-9.0210***	[-2.8]	-8.6037***	[-2.8]
γ_{nc}	-21.4486	[-0.4]	–	–
$\beta_{\varepsilon_{\overline{h}_{m,t}}}$	0.0584	[1.3]	0.0571	[1.3]
$\beta_{\varepsilon_{y_t^*}}$	0.0365	[1.5]	0.0385*	[1.7]
σ^2	78.2137***	[27.7]	78.2975***	[27.8]
log-likelihood	-6734.47		-6734.6	
observations	1906		1906	

NOTE. Excluded instruments for y_t: male and female age and age squared, education variables, male year of birth and male wages. Excluded instruments for $h_{m,t}$: male wage, male education variables. Excluded instruments for female wages: female age, female education variables, male job type. ***, **, * indicate significance at the 1, 5 and 10% level.

excluded instruments. Male and female age and age squared, male year of birth dummies, male education dummies and male wage rates are used to instrument y_t^*. Male wage rates and male education dummies are used as excluded instruments in the male labor supply equation. The residuals of both auxiliary regressions (the *estimated errors*) are included additively in the labor supply model. The null of exogeneity of both male market labor supply and full expenditure is tested with a t-test on the statistical significance of the two estimated error variables. The parameters associated with the estimated error variables are $\beta_{\varepsilon_{\overline{h}_{m,t}}}$ and $\beta_{y_t^*}$.

Table (5.1) reports regression results of two versions of equation (5.28). The first model allows for a flexible γ_{nc}. The second model restricts γ_{nc} to zero such that the model collapses to the linear expenditure system. The estimate of γ_{nc} from the first specification is not significantly different from zero. We are therefore unable to reject separable preferences over non-durable goods and female non-market time.[19]

---

[19] Note however that performing a t-test on significance of γ_{nc} is effectively performing a joint test on parameter significance and model specification. The inability to reject separability may be due of our specific parametric assumptions. Yet, the test clearly favors the second version of the model

The parameters of interest –the β's and the γ's– are not straightforward to interpret. We can conclude however that preferences for female non-market time change in the presence of children in the sense that time at home becomes more important. Focussing on the β parameters in table (5.1), we can conclude that when children grow older the effect of having children on preferences for non-market time becomes smaller. These results indicate that the extension of the traditional dichotomous tradeoff between leisure and consumption into a model that incorporates home production is important. Within our model, preferences for consumption and non-market time are in fact composites of preferences for consumption, leisure, domestic technology parameters and preferences for home produced goods (of which child care is an important one). From parameterizing preferences for non-market time with dummies of having children in certain age groups we therefore obtain a tentative estimate for preferences for child care. This under the assumption that households without children have no particular interest in child care. A final conclusion is that the β parameter increases with age and decreases in male hours of labor (significant impact of male hours on the female non-market hours - nondurable consumption tradeoff indicates non separabilities between these variables). The estimates associated with the estimated error variables are only significantly different from from zero at 10% for the full expenditure measure in the second column offering weak evidence for the econometric endogeneity of full expenditures. We do not find evidence for the econometric endogeneity of male labor supply.

5.4.2 The first stage: the Euler equation of non-durable consumption

So far we have estimated parameters of the within period model. The parameters of F however, capture the household's willingness to reallocate expenditures across time in response to financial or demographic incentives and have not yet been estimated. A convenient vehicle for estimating parameters of F are Euler equations of the equation (5.11d) type. Equation (5.11d) is one of the first order conditions of the consumer problem and relates marginal utility of wealth now and in the future. There are broadly three possibilities for operationalizing equation (5.11d) for

---

over the first one as it needs less parameters to fit the data.

estimation purposes:

$$\lambda_t = F_\Psi \Psi_{y_t^*} \tag{5.31}$$

$$\lambda_t = \frac{1}{w_{f,t}} \left(F_u u_{n_{ft}} - \mu_{1ft} \right) \tag{5.32}$$

$$\lambda_t = \frac{1}{p_t^{ND}} \left(F_u u_{c_t^{ND}} \right) \tag{5.33}$$

Combining equation (5.31) and equation (5.11d) yields an Euler equation of full expenditures. A combination of equation (5.32) and equation (5.11d) yields an Euler equation of female non-market time. A combination of equation (5.33) and equation (5.11d) yields an Euler equation of non-durable consumption.[20] Perhaps the most obvious parameterization of λ_t to use in estimation is equation (5.31), simply because we have specified Ψ in section (5.2.2). A functional form for the direct utility u has not been specified so far.

A key issue here is that when a household's optimal allocation involves corner solutions. We have argued in the previous section that corner solutions are particularly important for labor supply behavior (i.e., there are lot's of women who do not work). The formulas for the proposed Euler equations of non-market time and that of full expenditures –as a result– would contain the *unobserved* Kuhn-Tucker multipliers μ_{1ft} that are associated with the non-negativity constraints on female market labor supply. As a consequence, the indirect utility function (5.13) would no longer be a valid representation of preferences. If restrictions on the labor market are binding one should specify (indirect) utility functions of the restricted type (Neary and Roberts 1980).

Selection on working females only would be necessary to estimate Euler equations of full expenditure or female non-market time. Such a selection however, would be inappropriate for at least two reasons. First, because we intend to estimate Euler equations on the basis of constructed cohort averages we will introduce significant sampling error as the number of observations within each cohort will be small. A second issue is perhaps even more important. The selection on females with *current* positive labor supply introduces a selection bias. Euler equations predict that the error terms of the specified models are genuine forecast errors (mean independent of period t information). As a result, the errors will average out to zero when T gets large. The selection on workers has the important adverse side effect that the errors no longer average out to zero, because households with negative realizations of the forecast error will have a greater probability to be eliminated from the data set. However, the eliminated households should have been kept in the sample to

---

[20] Other, less intuitive combinations would be possible as well.

counterbalance the households who have received positive shocks.

We can safely assume that non-durable consumption is always positive and is therefore not in a corner solution. Consequently, the most profitable approach to estimate the parameters of F, that bypasses the trouble of non-zero Kuhn-Tucker multipliers or selection issues, is combining equation (5.33) with equation (5.11d) and construct an Euler equation of non-durable consumption. However, it is not evident to derive the Euler equation of non-durable consumption from an indirect utility function.

To derive Euler equations of non-durable consumption we use one of the empirical results from section (5.4.1). Because the estimate for γ_{nc} is insignificantly different from zero in the first column of table (5.1), we were not able to reject the linear expenditure system (LES). For the linear expenditure system there exists a direct (within period) utility function that corresponds to the indirect utility function we define in section (5.2.2). It is straightforward to derive Euler equations of non-durable consumption from direct utility functions.

The direct utility function corresponding to the LES is known as the Stone-Geary utility function. Equation (5.12) relates the Stone-Geary utility function to the indirect utility function specified in section (5.2.2) under the restriction that $\gamma_{nc} = 0$. The γ's and β's from the direct utility function correspond to the γ's and β's from the indirect utility function. In the remainder of this section we suppress the household i subscripts for clarity.

$$u\left(c_t^{ND}, n_{f,t}, z_{1t}\right) = \theta\left(n_{f,t} - \gamma_n\right)^{\beta_n}\left(c_t^{ND} - \gamma_c\right)^{(1-\beta_n)} \tag{5.34}$$

where $\theta = \beta_n^{-\beta_n}\left(1 - \beta_n\right)^{-(1-\beta_n)}$. We combine the above definition with condition (5.33) and (5.11d) of the household's optimization problem. We subsequently impose $\gamma_n^* = \gamma_n - T_f$ and $n_{f,t} = T_f - h_{f,t}$ to derive the (exact) Euler equation of non-durable consumption.

$$\begin{aligned}\frac{1+r_{t+1}}{1+\delta}\frac{\left(\frac{-h_{ft+1}-\gamma_n^*}{c_{t+1}-\gamma_c}\right)^{\beta_n(1-\rho)}\left(c_{t+1}-\gamma_c\right)^{-\rho}\theta\left(1-\beta_n\right)\times\exp\left[\alpha' z_{2t+1}\right]}{\left(\frac{-h_{ft}-\gamma_n^*}{c_t-\gamma_c}\right)^{\beta_n(1-\rho)}\left(c_t-\gamma_c\right)^{-\rho}\theta\left(1-\beta_n\right)\times\exp\left[\alpha' z_{2t}\right]}& \\ = 1+\varepsilon_{t+1}&\end{aligned} \tag{5.35}$$

where ε_{t+1} is a forecast error, such that its expected value is conditionally mean independent on information known at t:

$$E_t\varepsilon_{t+1} = 0 \tag{5.36}$$

Where E_t is the expectations operator, conditional on the information set of the decisionmaker at period t. The term $\theta\left(1-\beta_n\right)$ does not cancel out as β_n and hence θ are household and time specific by depending on z_{1t}.

We take *log*'s on both side of the Euler equation and apply a Taylor expansion of $[\ln 1+\varepsilon_{t+1}]$ around $\varepsilon_{it}=0$.

$$\begin{aligned}\Delta\ln\left(c_{t+1}-\gamma_c\right) &= \alpha_0+\frac{1}{\rho}\ln\left(1+r_{t+1}\right)+\frac{1-\rho}{\rho}\Delta\beta_n\ln\left(\frac{-h_{ft+1}-\gamma_n^*}{c_{t+1}-\gamma_c}\right)+\\ &\quad \frac{1}{\rho}\Delta\ln\left[\theta\left(1-\beta_n\right)\right]+\frac{1}{\rho}\alpha'\Delta z_{2t+1}-\\ &\quad \frac{1}{\rho}\left(\varepsilon_{t+1}+O_{t+1}\left(2\right)\right) \qquad (5.37)\end{aligned}$$

$\alpha_0=-\frac{1}{\rho}\ln\left(1+\delta\right)$ is constant and a function of the rate of time preference δ and the CRRA parameter ρ. The error term is written as the sum of the original forecast error ε_{t+1} and $O\left(2\right)$, which is a linear function of second and higher order moments of the forecast error.[21]

In estimation we assume that the r.h.s. composite variable $\frac{1}{\rho}\Delta\ln\left[\theta\left(1-\beta_n\right)\right]$ is sufficiently accounted for by a constant and the vector of taste shifters in first differences Δz_{2t+1}. Δz_{2t+1} contains changes is log family size, the number of children and a dummy for being single. The regression model becomes:

$$\begin{aligned}\Delta\ln\left(c_{t+1}-\gamma_c\right) &= \tilde{\alpha}_0+\frac{1}{\rho}\ln\left(1+r_{t+1}\right)+\frac{1-\rho}{\rho}\Delta\beta_n\ln\left(\frac{-h_{ft+1}-\gamma_n^*}{c_{t+1}-\gamma_c}\right)+\\ &\quad +\tilde{\alpha}_1\Delta\ln\mathrm{fs}_{t+1}+\tilde{\alpha}_2\Delta\mathrm{nch}_{t+1}+\tilde{\alpha}_3\Delta\mathrm{single}_{t+1}+\nu_{t+1} \qquad (5.38)\end{aligned}$$

The parameters γ_n^*, γ_c and β_n in (5.38) are replaced with their respective estimates [see table (5.1) column 2]. We treat the estimated parameters as known constants when estimating the other parameters of equation (5.38). We subsequently construct cohort means of the equation (5.38) variables using sampling weights. Our data set then covers 13 cohorts spanning 23 years (1978-2000) yielding 255 observations (some observations were missing at the beginning of the sample for the young cohorts and some were missing at the end for the older cohorts). We use the real interest rate (using yearly returns on government bonds as the nominal interest rate and the CPI as a price index. source: `http://statline.cbs.nl/`).

The parameter ρ appears twice in equation (5.38) and is therefore overidentified. We have tested this overidentifying restriction as a specification test rather than

---

[21] Using a Taylor expansion we can write: $\frac{1}{\rho}\ln\left(1+\varepsilon_{t+1}\right)=\frac{1}{\rho}\left(\varepsilon_{t+1}+O\left(2\right)\right)$. $O\left(2\right)$ is linear in the of higher order moments of the forecast error ε_{t+1}.

imposing the restriction at the outset. This is because hours of labor supply are predicted outside the 1988 to 1991 waves and because the consumption variables are typically measured with error. The composite variable $\Delta\beta_n \ln\left(\frac{-h_{ft+1}-\gamma_n^*}{c_{t+1}-\gamma_c}\right)$ therefore seems to be seriously affected by measurement error. We therefore tend to favor the estimates on ρ associated with the variable $\ln(1+r_{t+1})$. We use this estimate of ρ therefore for constructing elasticities in the subsequent section.

The life cycle hypothesis predicts $E_t\varepsilon_{t+1} = 0$ and not $E_t\nu_{t+1} = 0$. Methods of moments type estimators applied to equation (5.38) therefore, does not necessarily yield consistent estimators. The error term of the log-linearized Euler equation ν_{t+1} is a composite of first and higher order moments of the forecast error. For identification we therefore need that higher order moments of the forecast error ε_{t+1} are mean independent of period t variables as (i.e., $E_tO_{t+1}(2) = EO_{t+1}(2)$). $EO_{t+1}(2)$ will be consequently subsumed in the intercept.

Forcing this assumption on the data rules out potentially important phenomenon such as buffer stock savings behavior (Deaton 1991). It can be shown that for buffer stock consumers the conditional variance of the forecast error is a function of period t cash-on-hand (i.e., current income and assets) which invalidates $E_tO_{t+1}(2) = EO_{t+1}(2)$. If buffer stock behavior is important we could find that consumption growth is *excessively sensitive* to current income (at least if current income proxies for cash-on-hand). However, excess sensitivity of period $t+1$ consumption growth to period t income variables does not imply buffer stock saving behavior as binding borrowing constraints or specification errors such as intertemporal or within period non-separabilities might yield similar findings. All of these phenomenon are typically hard to disentangle empirically. We test the statistical importance either one (or a combination) of these mechanisms by including *log* period t real household income as an additional regressor in the Euler equation.

Innovations to the interest rate and the composite variable are directly associated with the forecast error and are therefore instrumented. The composite variable is also instrumented, because it is simultaneously determined with the dependent variable. Lagged interest rates (lagged one and two periods) are used to instrument the interest rate. One and two period lagged female labor supply variables and a two period lagged composite variable[22] and family composition dummies are used to instrument the composite variable. Households are assumed to perfectly anticipate changes in household composition one year in advance.

Table (5.2) column 1 reports the estimates the baseline regression of equation

---

[22] Due to possible measurement error in consumption, the composite error term in the log linearized Euler equation includes a MA(1) term such that one period lagged variables are bad instruments. In the presence of measurement error in the consumption variable, one period lagged variables are correlated with the error term. The interest rate is a macro variable and one period lags are therefore be assumed uncorrelated with the MA(1) term.

Table 5.2: *Estimation Results of the Euler Equation of Non-durable Consumption*

VARIABLE	(1)	(2)	(3)
r_{t+1}	0.053	0.138	0.227
	(0.298)	(0.320)	(0.345)
$\Delta\left[\beta_n \ln\left(\frac{-h_f-\gamma_n^*}{c-\gamma_c}\right)\right]_{t+1}$	-0.700***	-0.287	-0.114
	(0.186)	(0.262)	(0.273)
$\Delta \ln \mathrm{fs}_{t+1}$	0.794***	0.666***	0.614***
	(0.108)	(0.125)	(0.136)
Δ nch_{t+1}	-0.154***	-0.107**	-0.103**
	(0.041)	(0.048)	(0.051)
Δ single_{t+1}	0.025	-0.004	-0.055
	(0.094)	(0.106)	(0.122)
ln inc_t	–	–	-0.040**
			(0.020)
Constant	-0.011	0.051***	0.465**
	(0.014)	(0.018)	(0.214)
birth cohort dummies included	no	yes (p-val: 0.01)	yes (p-val: 0.00)
overid test on ρ	p-val: 0.47	p-val: 0.14	p-val: 0.15
lagged income included in instrument list	no	no	yes
Sargan test on the instruments	p-val: 0.04	p-val: 0.42	p-val: 0.40
Observations	229	229	229

NOTE. Robust standard errors in parentheses * significant at 10%; ** significant at 5%; *** significant at 1% $\Delta \ln \mathrm{fs}$ is the change in log family size. Δnch is the change in the number of children within the household. Δsingle is the change in the dummy for singles. The p-values associated with the birth cohort dummies test the hypothesis of excluding the dummies. The p-values associated with the overidentification test on ρ test the hypothesis whether the two estimates for ρ are the same [see equation (5.38)].

EXCLUDED INSTRUMENTS: Real interest rate lagged one and two periods, female hours of labor lagged one and two periods, the composite variable $\beta_n \ln\left(\frac{-h_f-\gamma_n^*}{c-\gamma_c}\right)$ lagged two periods and current period family composition dummies. The Sargan test on the validity of the over-identifying restrictions cannot be rejected for SPEC. 2 and SPEC. 3. There is evidence of excess sensitivity to current income variables.

(5.38). Table (5.2) column 2 adds birth cohort dummies to the regression (the youngest cohort is the baseline). Birth cohort dummies are important in explaining the growth rates of non-durable consumption where older cohorts experience slower consumption growth (i.e., younger cohorts have steeper consumption paths). Table (5.2) column 3 includes current log real household income as an additional regressor. The Sargan criterion indicates some model specification errors for column 1 which is resolved after including birth cohort dummies in column 2. The results of column 1 and (2) indicate important unobserved heterogeneities across cohorts. The parameter associated with the interest rate in column 2 is the estimate for $\frac{1}{\rho}$ that we use for calculating the elasticities in section (5.4.3). The estimate for ρ itself becomes $\hat{\rho} = 1/0.138 = 7.25$. Parameterization of $\frac{1}{\rho}$ by discriminating older and younger cohorts did not yield any significant differences.

It is somewhat unfortunate that the estimation results for the $1/\rho$ are not very precise. As a result, the elasticities that are calculated on the basis of this estimate are imprecise as well. This finding, however, is not uncommon in the literature [see e.g., Attanasio and Weber (1995) and Vissing-Jörgensen (2002)]. Departures from the standard log-linear approach into more advanced econometric techniques seem fruitful. Alan and Browning (2003)'s simulated residual estimation for example seems promising in reducing the size of standard errors. For the moment however, we conclude that households are not very much inclined to move expenditures back and forth through time in response to changes in financial incentives like the interest rate. Demographic variables, as measured by family size and the number of children, are much more important.

As a specification test we have included current real household income ($\ln inc_t$) as an additional regressor (table 5.2 column 3). We find that period t income is negatively and significantly related to consumption growth at $t + 1$, pointing to a rejection of one of the underlying assumptions of the life cycle model presented in this research, corroborating the findings of chapter 4. All else equal, low income cohorts experience high consumption growth rates on average.

Addressing the possible causes of excess sensitivity is an important issue for interpreting the elasticity estimates of following section. Buffer-stock behavior (Deaton 1991), borrowing constraints (Zeldes 1989) as well as intertemporal non-separabilities, or other specification errors could potentially produce these outcomes. It is beyond the scope of this study to determine the causes of this finding. It is important to realize however, that if intertemporal non-separabilities matter, the standard two-stage budgeting procedure may be invalidated as lagged expenditures would then (potentially) affect within period decisionmaking.

Under buffer stock behavior or under binding liquidity constraints the within period model is still consistently estimated. The significance of the income variable in table (5.2) column 3 therefore, does not imply a straight rejection of the life cycle model as a representation of preferences. The compensated and the uncompensated elasticities we report in the subsequent section may therefore be interpreted. Also, under buffer stock behavior or under (occasionally binding) borrowing constraints estimating intertemporal preference parameters is not necessarily hopeless as $1/\rho$

may be estimated with some degree of precision. If that is the case, the Frisch elasticities that we present in the following section can be interpreted, albeit with some degree of caution. For these households however, changes in financial incentives such as the interest rate are not that important as –because of their impatience– they are not very inclined to save anyway.[23]

The estimates of ρ do not change much under after inclusion of the income variable. If the buffer stock savings mechanism is important, income may be correlated with the higher order moments of the forecast error ε_{t+1} without invalidating life cycle theory in its current form. Theory however provides no clear argument why the interest rate –as one of the key instruments– and higher order moments of the forecast error would be correlated. Hence, offering an explanation of why the results in column 2 and (3) are rather similar. Other explanations for our findings may be equally plausible however. If intertemporal non-separabilities are important (e.g., habit formation) we should be cautious when interpreting our results. As a robustness check (not reported) we have included period t male market labor supply to see whether this had an effect on the significance of the income variable but it did not. We would like the reader to keep this in mind when interpreting the elasticities we provide in section (5.4.3).

5.4.3 Elasticities

With all relevant preference parameters now to our disposal we are able to calculate group specific elasticities. The formulas for the compensated, uncompensated, Frisch, and full expenditure elasticities are derived in appendix C.1. The elasticities are functions of the key variables we are studying. It makes sense therefore to construct interesting subgroups and see how elasticities differ across groups [see Blundell and Walker (1986) who do the same]. First of all, labor supply elasticities do not make much sense for non-workers. The elasticities presented in table (5.3) therefore are conditional on positive female hours of market labor supply.

It seems natural to use the data set we have available for constructing groups and simultaneously getting a feel of how important these groups are for the Dutch population. Within our sample for example, there are 444 households that exhibit the following features: no children in the household, the female in the household is under thirty years of age and the female is working. Because the total number of households in the data of which the female is working is 1906, the subgroup defined above represents approximately 23% ($= \frac{444}{1906} \times 100\%$) of the female work force in the

---

[23] The concept of impatience is important here. Patient households are typically inclined to save, rather than to borrow (against future labor income). For patient households therefore, liquidity constraints and buffer stock behavior is not an issue.

Netherlands.

Various groups may be discriminated on the basis of the observables within our data set. We have selected groups of households on the basis of female age and we discriminate households with and without children. Other selections would also be possible, but are not reported. Table (5.3) reports labor supply, non-durable consumption and full expenditure elasticities. We use the parameter estimates of column 2 of table (5.1) and the estimated parameter for $\rho = 7.25$ from column 2 of table (5.2). We subsequently construct average elasticities within each subgroup. We also report average age within each subgroup, average predicted hours of female market labor supply $\hat{h}_f$, average hours of female labor supply h_f, male hours of market labor supply within each subgroup h_m (this is important as the predicted hours as well as the elasticities are conditional on male hours of labor supply), and the number of observations within each subgroup n.

Section (5.2.1) argues that under specific assumptions on domestic technology, female non-market time may be regarded as a linear combination of hours of leisure and the home produced good (of which we consider child care as one of the most prominent types). Preferences for child care are strongly affected by the presence of children such that households substitute from market labor supply toward home production, both conditional on full expenditure y_t^* as conditional on the marginal utility of wealth λ_t.

Table 5.3: ELASTICITIES

							λ-constant elasticities				Full expenditure elast.			
Female age	kids	mean age	$\hat{h}_f$	h_f	h_m	n	E_{hh}	E_{hc}	E_{cc}	E_{ch}	η_h	η_h^*	η_c	η_c^*
<30	no	25.4	34.0	33.9	31.2	444	0.6	-0.5	-0.6	0.4	-1.1	-0.1	1.5	0.1
30-40	no	33.7	32.8	33.3	24.9	263	0.6	-0.5	-0.6	0.4	-1.2	-0.1	1.6	0.0
40-50	no	44.5	26.4	27.6	20.9	163	1.1	-0.8	-0.7	0.5	-1.9	-0.3	1.5	0.1
>50	no	54.4	23.8	22.3	15.7	175	1.2	-0.9	-0.7	0.5	-2.3	-0.3	1.5	0.1
<30	yes	27.2	17.1	16.3	33.7	101	2.7	-2.1	-0.9	0.7	-3.8	-1.0	1.5	0.3
30-40	yes	34.5	17.4	18.4	35.7	481	2.7	-2.1	-0.9	0.7	-3.7	-0.9	1.5	0.3
40-50	yes	42.7	17.7	19.4	35.4	268	2.6	-2.0	-0.9	0.7	-3.6	-0.9	1.4	0.3
>50	yes	51.1	12.2	13.7	14.6	11	7.8	-5.9	-0.9	0.7	-12.3	-3.9	1.4	0.3
total		35.8	24.7	25.2	29.8	1906	1.7	-1.3	-0.8	0.6	-2.5	-0.5	1.5	0.2

							Compensated elasticities				Uncompensated elasticities			
Female age	kids	mean age	$\hat{h}_f$	h_f	h_m	n	e_{hh}	e_{hc}	e_{cc}	e_{ch}	u_{hh}	u_{hc}	u_{cc}	u_{ch}
<30	no	25.4	34.0	33.9	31.2	444	0.6	-0.6	-0.5	0.5	0.2	-0.1	-1.1	1.0
30-40	no	33.7	32.8	33.3	24.9	263	0.6	-0.6	-0.5	0.5	0.2	-0.1	-1.1	1.1
40-50	no	44.5	26.4	27.6	20.9	163	0.9	-0.9	-0.6	0.6	0.5	-0.2	-1.1	1.0
>50	no	54.4	23.8	22.3	15.7	175	1.0	-1.0	-0.6	0.6	0.5	-0.2	-1.1	1.0
<30	yes	27.2	17.1	16.3	33.7	101	2.5	-2.5	-0.8	0.8	2.0	-1.0	-1.3	1.0
30-40	yes	34.5	17.4	18.4	35.7	481	2.4	-2.4	-0.8	0.8	1.9	-1.0	-1.4	1.1
40-50	yes	42.7	17.7	19.4	35.4	268	2.4	-2.4	-0.8	0.8	1.9	-0.9	-1.3	1.0
>50	yes	51.1	12.2	13.7	14.6	11	6.8	-6.8	-0.8	0.8	6.3	-2.4	-1.3	1.0
total		35.8	24.7	25.2	29.8	1906	1.5	-1.5	-0.6	0.6	1.1	-0.5	-1.2	1.0

NOTES: The subscripts $_{hh}$, $_{hc}$, $_{cc}$ and $_{ch}$ denote own price (= wage) elasticity of labor, cross price elasticity of labor, own price elasticity of consumption and cross price (= wage) elasticity of consumption respectively

Table (5.3) shows clear differences in predicted female hours of market labor supply $\hat{h}_f$ for households with and without children, after conditioning on age. When the average woman in the labor force has her first child, the model predicts that she will reduce her hours of labor from 30 to 18 hours. This drop should not be explained by a sudden change in preferences for leisure, but merely by an increase in preferences for child care (i.e., home production). Splitting up the sample in households with and without children again shows the significance of considering the effects of child care on female labor supply behavior. The reduction in hours after having children is easily explained within our model.

The average female labor supply uncompensated wage elasticity is about 1.1 which is somewhat larger than the elasticities that have been found in previous empirical studies on Dutch data (Theeuwes and Woittiez 1992, Evers et al. 2008). Discriminating (working) women with and without children reveals some interesting deviations from the global average. Working women without children work 30 hours per week on average, but seem hardly responsive to wage changes. The estimated uncompensated wage elasticity for this group is about 0.3. Working women with children work about 18 hours per week on average and are estimated to be much more responsive to wage changes, with an average elasticity of around 2. (Excluding the > 50 category with children does not affect the overall picture.)

Note that we have been estimating a model of female labor supply conditional on male hours of labor. Male hours are significant predictors for female hours and therefore affect the elasticities that are reported in the table. Conditional on positive female hours, males work 35 hours on average when there are children. When there are no children, males work only 25 hours on average.

The Frisch elasticities take the intertemporal re-allocation of full expenditures into account and reveal some additional interesting patterns. We find that ρ is imprecisely estimated and seems relatively large compared to what has been found in other studies. Moreover, we find evidence against the life cycle hypothesis as period t current household income is a significant predictor of consumption growth at $t + 1$. On the whole we need to be careful with interpreting Frisch elasticities. The large ρ indicates that households are not quite willing to reallocate resources over time in response to financial incentives. From the formulas of the Frisch elasticities [equation (C.22), (C.24), (C.26) and (C.28)] we can infer that for increasing ρ Frisch elasticities converge to the compensated elasticities [equation (C.12), (C.14), (C.16) and (C.18)]. The average Frisch female labor supply elasticities across household categories is about 1.7. The Frisch elasticities on average about double the size of the un-compensated elasticities for the households without children. Frisch elasticities for households that have children are about 30 percent larger than the uncompensated elasticities. The intertemporal reallocation of expenditures seems to play a more important role in allocating time when there are no children in the household. Recalculating of the elasticities with a smaller value for ρ imputed [we have used $\rho = 3$] increases the Frisch elasticities in absolute value by about 20%.

We also wish to emphasize that the Frisch elasticities estimate the response to a *temporary* price or wage change. E_{hh} for example, predicts that labor supply responses only to a *current* wage increases while (expected) future wages are unchanged. A larger set of questions could analyzed within a simulation study. Not only could we predict current marginal wage/price effects, but effects of changes in expected income patterns or even changes in uncertainty about these income patterns. Finally, the η^* and η represent full expenditure elasticities on the basis of y_t^* and y_t respectively. For constructing η we fixed the total time endowment at $T_f = 90$.

5.5 Conclusion

This research uses the expected, exponentially discounted utility model as the organizing framework to analyze the allocation of female time and non-durable consumption over the life cycle. We condition on a number of household demographic characteristics and male labor supply. We summarize the principal conclusions of this chapter:

We show that when the production technology for non-marketable home produced goods exhibits constant returns to scale and when the factors of production are perfect substitutes, standard empirical models of (female) labor supply may be consistently estimated and interpreted without imposing undesirable separability assumptions on the home produced good. We show that preference parameters for consumption and non-market time are functions of preferences associated with consumption, leisure, home production and production technology parameters. This fact calls for a structural reinterpretation of parameter estimates of standard labor supply models and points to the importance of allowing for demographic characteristics in estimation. The sharp decrease in average predicted hours of labor when women are having their first child indicates the need for this change of focus.

Working women without children work 30 hours on average as opposed to the 18 hours of average market labor supply of working women with children. Working women with children are however much more responsive to wage changes (with an elasticity of about 2) than women without children (with an elasticity of about 0.3). We show that when optimal intertemporal allocation of resources is taken into account, wage elasticities of female labor supply increase by about 50% on average (from 1.1 to 1.7).

We have estimated an Euler equation of non-durable consumption using a large

T micro data set (we have constructed a synthetic panel spanning 23 years). The Euler equation is employed for identifying intertemporal preference parameters (the parameters of F in this study). The parameters of intertemporal allocation are rather small and imprecisely estimated indicating that Dutch households are not particularly responsive (in terms of adjusting their growth path) to changes in financial incentive such as the interest rate.

We find evidence against the life cycle hypothesis by testing for excess sensitivity of expected consumption growth to real current household income. We find this result after conditioning on family size and number of children (and male labor supply in a robustness check). There are multiple explanations for this finding ranging from speciation errors, such as habit formation, to other concepts like buffer stock savings behavior or binding liquidity constraints. Another explanation may be that households behave more ad hoc, and consume just a fraction of current income. Disentangling these concepts empirically is typically hard and is beyond the scope of this study.

Chapter 6

Conclusions

I start the conclusion of this thesis with two remarks: *1.* subjective data or self-reports, can be a useful source of information to infer important elements of individual preferences that would otherwise go undiscovered. Great progress can be attained by further developing theories that explain observed actions and e.g., self-reports on satisfaction in one integrated framework. *2.* I would like to express a degree of discomfort with the structural modeling approach that is widely adopted in applied microeconometrics (and, I admit, my thesis is no exception). If theoretical models or statistical assumptions are heavily debated, the benefits of estimating so-called structural parameters is questionable, because the parameters often have no clear interpretation if the theory is false. In such situations it is better to use theories as *guidelines* to understand real-world phenomena, rather than to superimpose them on reality as if they were exactly true.

1. In the introduction I have already discussed the potential of self reported or stated preferences data. In general, I argue, subjective data can be used to address problems of identification that can be encountered in revealed preferences analysis [see chapter 2 and chapter chapter 3 of this thesis]. The identification problems that can be encountered include: limited possibility of action (constraints on the labor or capital market), preferences that do not necessarily lead to action (satisfaction, opinions), and in situations where normative preferences and behavior (i.e., action) do not match (e.g., naive hyperbolic discounting, complexity, myopia, etc. see Beshears et al. (2008) for an interesting exposition about this topic).

Whereas there seem ample possibilities for using subjective data in economics, there is, as for now, no generally accepted theory on how to interpret some of this data [see Kahneman et al. (1997) for developments in this direction]. I would like to focus for a moment on a widely analyzed and interesting type of stated preferences data: satisfaction data. Many surveys include questions on household or individual satisfaction with various elements of personal life, from satisfaction with income, health status, to political regimes (e.g., see the German Socio-Economic Panel `http://panel.gsoep.de/`). Whereas analyzing the properties of these variables is interesting in its own right, progress may be attained by developing theories that simultaneously explain both satisfaction data and behavioral outcomes. For example, I find it intuitive that income satisfaction and labor supply are results of the same underlying (maximization) process. I would like to hypothesize that income satisfaction relates inversely to the *marginal* utility of income (i.e., high satisfaction with x means low marginal utility of x) [see also chapter 3]. Marginal utilities in turn, are useful concepts in economics. If one is willing to accept this interpretation, satisfaction data can be used to test, validate or to extend existing theories of behavior.

2. Structural modeling is widespread in applied microeconometrics. Structural modeling is what I have termed "revealed preferences analysis in practice" in the introduction of this thesis. A preference function is specified and given some statistical assumptions, "deep" or "structural" parameters of the preference function can be estimated using the right data. The idea is that these structural parameters are independent of context and consequently more useful for policy evaluations (i.e., the famous *Lucas critique*).

A disadvantage of this approach is that in many instances, the validity of the preference function as well as the statistical assumptions needed to estimate the parameters are "hard to defend in practice (Deaton (1997), page 3)". Summarizing the data by means of a set of estimated "structural" parameters is interesting, but *only* if the model and the statistical assumptions are approximately correctly specified. Problems arise if the framework is flawed, such that, as a result, the estimated structural parameters lose their original interpretation. Often, the transparency of econometric methodology is sacrificed in an attempt to do justice to the complexity of the process that is studied. My conception of this matter relates very much to the arguments put forward by Deaton (1997):"But in the end, I believe that we make more progress, not by pretending to estimate structural parameters, but by asking whether our theories and their policy implications are consistent with well-chosen stylized facts (Deaton (1997), page 4)." In situations where the validity of an economic model is (highly) questionable, applied researchers should focus on reporting properties of the data that do not depend on the theory in order to be meaningfully communicated.

Chapter 4 and chapter 5 of this thesis give a subtle distinction between *strict* structural modeling (in chapter 5) and a more *loose* approach (in chapter 4). The parameters, and the elasticities presented in chapter 5 must be interpreted within the context of the life cycle model that is specified. A rejection of the overidentifying restrictions that are implied by the Euler equation already puts the results into perspective. Furthermore, the supposition that durable consumption plays no role in the dynamic tradoff between nondurable consumption and female time and that (female) wages are exogenous with respect to labor supply, are perhaps too strong assumptions. These results therefore need to be interpreted with caution. Clearly, no transparent interpretation of the data can be communicated if the model and the statistical assumptions are wrong. In contrast, chapter 4 simply reports life cycle profiles of a few important variables. The methodology is clear, and the results can be explained ad hoc: household consumption increases until age 45 and decreases thereafter. These stylized facts can be easily communicated and are interesting. Moreover, these facts also link, in a much looser way, to a theoretical predictions of the expected utility/life cycle model.

Appendix A

Appendix: chapter 2

A.1 The conditional density function

$h\left(\mathbf{w}\middle|a,\mathbf{x},\mathbf{z},\mathrm{p};\theta^d,\sigma_e^2\right)$ is defined as follows:

$$h\left(\mathbf{w}\middle|a,\mathbf{x},\mathbf{z},\mathrm{p};\theta^d,\sigma_e^2\right)=\frac{1}{\sigma_e}\phi\left(\frac{w_1-w_1\left(\mathbf{x},\mathbf{z},\mathrm{p},\theta^d\right)-a}{\sigma_e}\right)$$

where $\phi\left(\cdot\right)$ is a standard normal p.d.f.

Function $g\left(A\middle|\eta,\mathbf{x},\mathbf{z},\mathrm{p},\theta^d,\theta^s\right)$ is constructed out of equation (2.11) and (2.17) and is defined as:

$$g=\left[P\left(A=1\middle|\mathbf{x},\mathbf{z},\mathrm{p},\theta^d,\theta^s,\eta\right)\right]^{1(A=1)}\times\left[P\left(A=0\middle|\mathbf{x},\mathbf{z},\mathrm{p},\theta^d,\theta^s,\eta\right)\right]^{1(A=0)}$$

where:

$$\begin{aligned}P\left(A=1\middle|\mathbf{x},\mathbf{z},\mathrm{p},\theta^d,\theta^s,\eta\right)&=\Phi\left(\frac{\ln\mathbf{x}-\ln a\left(\mathrm{p},\mathbf{z}\right)}{b\left(\mathrm{p},\mathbf{z}\right)}+\delta\left(\mathbf{z}\right)+\eta\right)\\P\left(A=0\middle|\mathbf{x},\mathbf{z},\mathrm{p},\theta^d,\theta^s,\eta\right)&=1-\Phi\left(\frac{\ln\mathbf{x}-\ln a\left(\mathrm{p},\mathbf{z}\right)}{b\left(\mathrm{p},\mathbf{z}\right)}+\delta\left(\mathbf{z}\right)+\eta\right)\end{aligned}$$

where Φ is the c.d.f. of the standard normal distribution.

A.2 Estimating non-food prices

Non-food prices in Indonesia (as measures for making purchasing power comparisons between regions and time) are not recorded. We follow Ravallion and Bidani (1994) by estimating the non-food prices. We stay close to the original (demand) system equation (2.21) and estimate non-food prices conditional on AIDS preferences. By imposing $\gamma_{11} = 0$ we can write the budget share equation as follows:

$$w_1 = (\alpha_1 - \beta_1\alpha_0) - \beta_1\alpha_1 \ln k_1 \mathrm{p}_1^r + \beta_1\alpha_1 \ln k_2 \mathrm{p}_2^r + \beta_1 \ln \frac{\mathbf{x}}{k_2 \mathrm{p}_2^r}$$

where p_2^r is the food poverty line. The p_1^r's are region specific non-food price indices that are estimated using dummy variables. This approach assumes that regional differences in consumption patterns that remain after conditioning on food prices, income levels and demographic variables, are due to differences in non-food prices. The working of the conditionality is subsequently described by the AIDS preferences (which is not unreasonable given figure (2.1)).

We interact α_0 with the *log* of household size and re-parameterize:

$$\begin{aligned} w_1 &= (\alpha_1 - \beta_1\alpha_0(\mathbf{z}) + \beta_1\alpha_1 - \beta_1 \ln k_2) - \beta_1\alpha_1 \ln k_1 \mathrm{p}_1^r + \beta_1\alpha_1 \ln \mathrm{p}_2^r + \beta_1 \ln \frac{\mathbf{x}}{\mathrm{p}_2^r} \\ &= \phi_0^0 + \phi_0^{\mathbf{z}} \ln fs - \phi_1 \ln k_1 \mathrm{p}_1^r + \phi_1 \ln \mathrm{p}_2^r + \phi_2 \ln \frac{\mathbf{x}}{\mathrm{p}_2^r} \end{aligned}$$

We are not interested in the arbitrary scaling parameters k_1 and k_2 nor in estimating the structural parameters. Impose $k_1^1 = 1/\mathrm{p}_1^1$. i.e., the first regions price is normalized to 1. For estimation we replace the p_1 term with region and time dummies.

$$w_1 = \phi_0^0 + \phi_0^{\mathbf{z}} \ln fs - \phi_1 \times \sum_{r=2}^{R} \gamma^r D^r + \phi_1 \ln \mathrm{p}_2^r + \phi_2 \ln \frac{\mathbf{x}}{\mathrm{p}_2^r}$$

re-parameterizing $\tilde{\gamma}^r = -\phi_1 \times \gamma^r$ yields a model that can be estimated with linear regression techniques:

$$w_1 = \phi_0^0 + \phi_0^{\mathbf{z}} \ln fs + \sum_{r=2}^{R} \tilde{\gamma}^r D^r + \phi_1 \ln \mathrm{p}_2^r + \phi_2 \ln \frac{\mathbf{x}}{\mathrm{p}_2^r}$$

Dividing $\tilde{\gamma}^r$ by $-\phi_1$ and constructing the exponent yields an estimate of non-food price index: $\mathrm{p}_1^r = \exp(\gamma^r) = \exp\left(-\hat{\tilde{\gamma}}^r / \hat{\phi}_1\right)$

A.3 Prices and parameterizations

It is widespread in empirical demand analysis to study composite goods (e.g., food and non-food consumption in this paper) using price indices rather than unit prices. In other words, prices are normalized in some way. An issue with using price indices rather than unit prices is that the normalization is directly affecting the size and significance levels of some, but not all, of the parameters of the model. It can be shown that the magnitude, sign and significance levels the $\beta_0^{\mathbf{z}}$ parameter, for example, depends on price normalizations in an arbitrary way (i.e., measuring food consumption in ounces or in kilos matters for the estimates for $\beta_0^{\mathbf{z}}$). This is another way of saying that $\beta_0^{\mathbf{z}}$ is not uniquely identified. Testing parameter restrictions or otherwise interpreting $\beta_0^{\mathbf{z}}$ therefore, is meaningless. We show here that for that reason, preference parameters must be appropriately interacted with demographic variables [e.g., one cannot parameterize β_1, without parameterizing β_0 in exactly the same way]. Otherwise, this inherit arbitrariness will seep through in our estimates of the equivalence scales. More generally: we typically do not want that price normalization matters for final outcomes (e.g., likelihood outcomes, elasticities, predictions or equivalence scales). It would not make sense if final outcomes depend on arbitrary linear scaling of the price vectors. This requirement affects ways in which can parameterize the structural parameters.

Note that prices only affect final outcomes through its effects on Ψ, here using the A.I.D.S. parameterization defined by (2.17), (2.18) and (2.19). To show the importance of being careful when interacting the structural parameters with a vector of demographics we are considering the model after renormalizing the price vector: $\mathrm{p}_1 = k_1\tilde{\mathrm{p}}_1$ $\mathrm{p}_2 = k_2\tilde{\mathrm{p}}_2$. This renormalization of the price indices would represent for example that we no longer consider food consumption in kilos but in ounces. This should not affect the outcomes in any fundamental way. We rewrite equation (2.18) and (2.19) after substituting the new price indices:

$$\begin{aligned} \ln a\left(\mathrm{p},\mathbf{z}\right) &= \ln a\left(k\tilde{\mathrm{p}},\mathbf{z}\right) \qquad\qquad (A.1)\\ &= \left[\alpha_0 + \alpha_1 \ln k_1 + (1-\alpha_1)\ln k_2 + 0.5\gamma_{11}\left(\ln\frac{k_1}{k_2}\right)^2\right] + \\ &\quad \left(\alpha_1 + \gamma_{11}\ln\frac{k_1}{k_2}\right)\ln\tilde{\mathrm{p}}_1 + \left(1-\alpha_1-\gamma_{11}\ln\frac{k_1}{k_2}\right)\ln\tilde{\mathrm{p}}_2 + 0.5\gamma_{11}\left(\ln\frac{\tilde{p}_1}{\tilde{p}_2}\right)^2 \\ b\left(\mathrm{p},\mathbf{z}\right) &= b\left(k\tilde{\mathrm{p}},\mathbf{z}\right) \qquad\qquad (A.2)\\ &= \exp\left[\beta_0 + \beta_1\ln\frac{k_1}{k_2}\right]\left(\frac{\tilde{\mathrm{p}}_1}{\tilde{\mathrm{p}}_2}\right)^{\beta_1} \end{aligned}$$

Both $\ln a\left(k\tilde{\mathrm{p}},\mathbf{z}\right)$ and $b\left(k\tilde{\mathrm{p}},\mathbf{z}\right)$ can be written such that the original structure of the

AIDS model is revealed:

$$\begin{aligned} \ln a\left(k\tilde{\mathrm{p}},\mathbf{z}\right) &= \ln\tilde{a}\left(\tilde{\mathrm{p}},\mathbf{z}\right) \\ &= \tilde{\alpha}_0+\tilde{\alpha}_1\ln\tilde{\mathrm{p}}_1+\left(1-\tilde{\alpha}_1\right)\ln\mathrm{p}_2+0.5\gamma_{11}\left(\ln\frac{\tilde{\mathrm{p}}_1}{\tilde{\mathrm{p}}_2}\right)^2 \end{aligned} \tag{A.3}$$

where:

$$\tilde{\alpha}_0 = \alpha_0+\alpha_1\ln k_1+\left(1-\alpha_1\right)\ln k_2+0.5\gamma_{11}\left(\ln\frac{k_1}{k_2}\right)^2 \tag{A.4}$$

$$\tilde{\alpha}_1 = \alpha_1+\gamma_{11}\ln\frac{k_1}{k_2} \tag{A.5}$$

Moreover:

$$\begin{aligned} b\left(k\tilde{\mathrm{p}},\mathbf{z}\right) &= \tilde{b}\left(\tilde{\mathrm{p}},\mathbf{z}\right) \\ &= \exp\left[\tilde{\beta}_0\right]\left(\frac{\tilde{\mathrm{p}}_1}{\tilde{\mathrm{p}}_2}\right)^{\tilde{\beta}_1} \end{aligned} \tag{A.6}$$

where:

$$\tilde{\beta}_1 = \beta_1 \tag{A.7}$$

$$\tilde{\beta}_0 = \beta_0+\beta_1\ln\frac{k_1}{k_2} \tag{A.8}$$

From equation (A.7) we can conclude that linear price transformations does not matter for $\tilde{\beta}_1$. Instead, the parameters $\tilde{\alpha}_0$, $\tilde{\alpha}_1$ and $\tilde{\beta}_0$ are depend on the k's and hence on linear price transformations.

The possibility to recover the original structure of the AIDS model after a transformation of prices proves that final outcomes do not depend on these price transformations. So, if meat consumption is measured in kilos rather than ounces (the unit of measurement is changed) we *will* get different parameter estimates, because the k parameters change. The predicted consumption of meat changes accordingly, simply because now we measure consumption in ounces rather than in kilograms. Predicted consumption of the other goods will not be affected. Obviously, a rescaling of prices is not affecting the *actual amounts* that are consumed. That is, price normalization does not matter for the fundamental predictions of the model. The fact that k_1 and k_2 are estimated alongside the other parameters account for this (in fact we are estimating composite parameters). We want this property to carry over to the case where we interact structural parameter with a vector of demographics $\mathbf{z}$.

If we are interested in differentiating consumption patterns of households of different sizes and compositions, inappropriately parameterizing the model parameters yield predictions that are arbitrary and hard to interpret. In everyday terminology: if an empirical study concludes that large families consume more meat (ceteris paribus) than small families, we want this outcome not to depend on whether meat consumption is measured in ounces or in kilograms.

For the sake of the argument we limit ourself to linear parameterizations of the structural parameters of the type $\alpha_0(\mathbf{z}) = \alpha_0^0 + \alpha_0^{\mathbf{z}}\mathbf{z}$ (non-linear interactions are rarely seen in the literature and indeed greatly complicate the matter). From (A.4) and (A.5) it can be readily inferred that we can parameterize α_0 without having α_1 affected. A linear parameterization of α_1 however requires that α_0 should be parameterized in the same way. If we parameterize α_1 *but not* α_0 we are basically imposing arbitrary assumptions on k, where in fact they should be estimated together with the model parameters. A similar story goes for the β parameters, where parameterizing β_1 requires that β_0 is parameterized in exactly the same way. If we do not commit to this modeling requirement we could find that differences in meat consumption across households of different sizes depend on whether meat is measured in ounces or in kilograms.

Appendix B

Appendix: chapter 3

B.1 Empirical operationalization

The expectation of the MRS_{t+1} ($= \frac{u'_{t+1}}{u'_t}$) can be written as follows:

$$E_t\left[MRS_{t+1}\right] = \int_{-\infty}^{\infty} f\left(MRS_{t+1}\middle|t\right) MRS_{t+1} dMRS_{t+1} \tag{B.1}$$

The conditional pdf $f\left(MRS_{t+1}\middle|t\right)$ can rewritten:

$$f\left(MRS_{t+1}\middle|t\right) = \sum_j f\left(MRS_{t+1}\middle|t, R_{t+1} = j\right) \times P\left(R_{t+1} = j\middle|t\right) \tag{B.2}$$

Where R_{t+1} is a discrete random variable that can take on the values $j = 1, 2, 3, 4, 5$ denoting the recollected changes in adequacy.

Equation B.1 and B.2 can be combined:

$$\begin{aligned} & E_t\left[MRS_{t+1}\right] \\ = & \int_{-\infty}^{\infty} \left(\sum_j f\left(MRS_{t+1}\middle|t, R_{t+1} = j\right) \times P\left(R_{t+1} = j\middle|t\right) \right) MRS_{t+1} dMRS_{t+1} \\ = & \sum_j P\left(R_{t+1} = j\middle|t\right) \int_{-\infty}^{\infty} f\left(MRS_{t+1}\middle|t, R_{t+1}\right) MRS_{t+1} dMRS_{t+1} \\ = & \sum_j P\left(R_{t+1} = j\middle|t\right) E\left[MRS_{t+1}\middle|t, R_{t+1}\right] \end{aligned} \tag{B.3}$$

B.2 Hayashi's proof

Under the assumption that $r_\tau = r \forall \tau$ rewrite equation (3.6) as follows:

$$E_t \left[\left((1+r) \beta \tilde{u}'_{t+1} (\cdot) - \tilde{u}'_t (\cdot) \right) - \alpha \beta \left((1+r) \beta \tilde{u}'_{t+2} (\cdot) - \tilde{u}'_{t+1} (\cdot) \right) \right] = 0 \qquad \text{(B.4)}$$

For expositional purposes Hayashi (1985) introduces:

$$E_t y_{t+k} = E_t \left[\left((1+r) \beta \tilde{u}'_{t+k+1} (\cdot) - \tilde{u}'_{t+k} (\cdot) \right) \right] \qquad \text{(B.5)}$$

such that equation (B.4) is written as:

$$E_t y_t - \alpha \beta E_t y_{t+1} = 0 \qquad \text{(B.6)}$$

Equation (B.7) is supposed to hold throughout life:

$$E_s y_s - \alpha \beta E_s y_{s+1} = 0, \quad s = t, t+1, t+2, \ldots, T-1 \qquad \text{(B.7)}$$

Taking the expectation of the above conditional on current period information:

$$E_t y_s - \alpha \beta E_t y_{s+1} = 0, \quad s = t, t+1, t+2, \ldots, T-1 \qquad \text{(B.8)}$$

Equation (B.8) is a first order difference equation and can be solved by backward induction:

$$E_t y_T = \left(\frac{1}{\alpha \beta} \right)^T E_t y_t \qquad \text{(B.9)}$$

The transversality condition prescribes that marginal utilities are not allowed to tend to infinity ("the terminal value is given" (Hayashi 1985)). Under $-1 < \alpha < 1$ and $0 \leq \beta < 1$ we must therefore conclude that $E_t y_t \to 0$ if $T \to \infty$ [this is the proof of Hayashi (1985)]. As a result we can write equation (B.4) as:

$$(1+r) \beta E_t \left[\frac{\tilde{u}'_{t+1} (\cdot)}{\tilde{u}'_t (\cdot)} \right] = 1 \qquad \text{(B.10)}$$

Appendix C

Appendix: chapter 5

C.1 Elasticities

For the sake of convenience we write the elasticities in terms of the *observable* y_t^*.

Uncompensated demands:

$$h_{f,t}^{\text{uncomp.}} = T - n_{f,t} \quad = \quad -a^*_{w_{f,t}} - \frac{b_{w_{f,t}}}{b}\left[y_t^* - a^*\right] \tag{C.1}$$

$$c_t^{\text{uncomp.}} \quad = \quad a^*_{p_t} + \frac{b_{p_t}}{b}\left[y_t^* - a^*\right] \tag{C.2}$$

uncompensated labor supply elasticities:

$$\frac{\partial h_{f,t}}{\partial w_{f,t}}\frac{w_{f,t}}{h_{f,t}} \quad = \quad \left[-a^*_{ww} - \frac{b_{ww}b - \left[b_w\right]^2}{b^2}\left[y^* - a^*\right] - \frac{b_w}{b}\left[-a^*_w\right]\right] \cdot \frac{w}{h} \tag{C.3}$$

$$\frac{\partial h_{f,t}}{\partial p_t}\frac{p_t}{h_{f,t}} \quad = \quad \left[-a^*_{wp} - \frac{b_{wp}b - b_w b_p}{b^2}\left[y^* - a^*\right] - \frac{b_w}{b}\left[-a^*_p\right]\right] \cdot \frac{p}{h} \tag{C.4}$$

$$\frac{\partial h_{f,t}}{\partial y_t^*}\frac{y_t^*}{h_{f,t}} \quad = \quad -\frac{b_w}{b} \cdot \frac{y^*}{h} \tag{C.5}$$

uncompensated consumption elasticities:

$$\frac{\partial c_t}{\partial w_{f,t}}\frac{w_{f,t}}{c_t} \quad = \quad \left[a^*_{pw} + \frac{b_{pw}b - b_p b_w}{b^2}\left[y^* - a^*\right] + \frac{b_p}{b}\left[-a^*_w\right]\right] \cdot \frac{w}{c} \tag{C.6}$$

$$\frac{\partial c_t}{\partial p_t}\frac{p_t}{c_t} \quad = \quad \left[a^*_{pp} + \frac{b_{pp}b - \left[b_p\right]^2}{b^2}\left[y^* - a^*\right] + \frac{b_p}{b}\left[-a^*_p\right]\right] \cdot \frac{p}{c} \tag{C.7}$$

$$\frac{\partial h_{f,t}}{\partial y_t^*}\frac{y_t^*}{h_{f,t}} \quad = \quad \frac{b_p}{b} \cdot \frac{y^*}{c} \tag{C.8}$$

$\overline{U}$ λ may be written in terms of y^*. $\overline{U} = \frac{y_t^*-a^*}{b}$ and $\lambda = \left(\frac{y_t^*-a^*}{b}\right)^{-\rho} \frac{1}{b} \times \exp\left[\alpha' z_{2t}\right]$.

Compensated demands:

$$\begin{aligned} h_{f,t}^{\text{uncomp.}} &= -a^*_{w_{f,t}} - b_{w_{f,t}}\overline{U} && \text{(C.9)} \\ h_{f,t}^{\text{uncomp.}} &= a^*_{p_t} + b_{p_t}\overline{U} && \text{(C.10)} \end{aligned}$$

such that:

$$\begin{aligned} \frac{\partial h_{f,t}}{\partial w_{f,t}}\frac{w_{f,t}}{h_{f,t}} &= \left[-a^*_{ww} - b_{ww}\overline{U}\right] \cdot \frac{w_{f,t}}{h_{f,t}} && \text{(C.11)} \\ &= \left[-a^*_{ww} - \frac{b_{ww}}{b}\left[y^* - a^*\right]\right] \cdot \frac{w_{f,t}}{h_{f,t}} && \text{(C.12)} \\ \frac{\partial h_{f,t}}{\partial p_t}\frac{p_t}{h_{f,t}} &= \left[-a^*_{wp} - b_{wp}\overline{U}\right] \cdot \frac{p_t}{h_{f,t}} && \text{(C.13)} \\ &= \left[-a^*_{wp} - \frac{b_{wp}}{b}\left[y^* - a^*\right]\right] \cdot \frac{p_t}{h_{f,t}} && \text{(C.14)} \end{aligned}$$

and:

$$\begin{aligned} \frac{\partial c_t}{\partial w_{f,t}}\frac{w_{f,t}}{c_t} &= \left[a^*_{pw} + b_{pw}\overline{U}\right] \cdot \frac{w_{f,t}}{c_t} && \text{(C.15)} \\ &= \left[a^*_{pw} + \frac{b_{pw}}{b}\left[y^* - a^*\right]\right] \cdot \frac{w_{f,t}}{c_t} && \text{(C.16)} \\ \frac{\partial c_t}{\partial p_t}\frac{p_t}{c_t} &= \left[a^*_{pp} + b_{pp}\overline{U}\right] \cdot \frac{p_t}{c_t} && \text{(C.17)} \\ &= \left[a^*_{pp} + \frac{b_{pp}}{b}\left[y^* - a^*\right]\right] \cdot \frac{p_t}{c_t} && \text{(C.18)} \end{aligned}$$

Frisch demand functions are:

$$\begin{aligned} h_{f,t}^{\text{Frisch.}} &= -a^*_{w_{f,t}} - \frac{b_{w_{f,t}}}{(b\lambda_t)^{\frac{1}{\rho}}} \times \exp\left[\frac{1}{\rho}\alpha' z_{2t}\right] && \text{(C.19)} \\ c_t^{\text{Frisch.}} &= a^*_{p_t} + \frac{b_{p_t}}{(b\lambda_t)^{\frac{1}{\rho}}} \times \exp\left[\frac{1}{\rho}\alpha' z_{2t}\right] && \text{(C.20)} \end{aligned}$$

such that:

$$\begin{aligned}
\frac{\partial h_{f,t}}{\partial w_{f,t}}\frac{w_{f,t}}{h_{f,t}} &= \left[-a^*_{ww} - \frac{b_{ww} - [b_w]^2(\rho b)^{-1}}{(b\lambda)^{\frac{1}{\rho}}}\exp\left[\frac{1}{\rho}\alpha' z\right]\right]\cdot\frac{w_{f,t}}{h_{f,t}} && \text{(C.21)}\\
&= \left[-a^*_{ww} - \frac{b_{ww} - [b_w]^2(\rho b)^{-1}}{b}\left[y^* - a^*\right]\right]\cdot\frac{w_{f,t}}{h_{f,t}} && \text{(C.22)}\\
\frac{\partial h_{f,t}}{\partial p_t}\frac{p_t}{h_{f,t}} &= \left[-a^*_{wp} - \frac{b_{wp} - b_w b_p(\rho b)^{-1}}{(b\lambda)^{\frac{1}{\rho}}}\exp\left[\frac{1}{\rho}\alpha' z\right]\right]\cdot\frac{p_t}{h_{f,t}} && \text{(C.23)}\\
&= \left[-a^*_{wp} - \frac{b_{wp} - b_w b_p(\rho b)^{-1}}{b}\left[y^* - a^*\right]\right]\cdot\frac{p_t}{h_{f,t}} && \text{(C.24)}
\end{aligned}$$

and:

$$\begin{aligned}
\frac{\partial c_t}{\partial w_{f,t}}\frac{w_{f,t}}{c_t} &= \left[a^*_{pw} + \frac{b_{pw} - b_p b_w(\rho b)^{-1}}{(b\lambda)^{\frac{1}{\rho}}}\exp\left[\frac{1}{\rho}\alpha' z\right]\right]\cdot\frac{w_{f,t}}{c_t} && \text{(C.25)}\\
&= \left[a^*_{pw} + \frac{b_{pw} - b_p b_w(\rho b)^{-1}}{b}\left[y^* - a^*\right]\right]\cdot\frac{w_{f,t}}{c_t} && \text{(C.26)}\\
\frac{\partial c_t}{\partial p_t}\frac{p_t}{c_t} &= \left[a^*_{pp} + \frac{b_{pp} - [b_p]^2(\rho b)^{-1}}{(b\lambda)^{\frac{1}{\rho}}}\exp\left[\frac{1}{\rho}\alpha' z\right]\right]\cdot\frac{p_t}{c_t} && \text{(C.27)}\\
&= \left[a^*_{pp} + \frac{b_{pp} - [b_p]^2(\rho b)^{-1}}{b}\left[y^* - a^*\right]\right]\cdot\frac{p_t}{c_t} && \text{(C.28)}
\end{aligned}$$

We normalize all the price data in the process such that non-durable prices become 1 and nominal wages become real wages.

Nederlandse samenvatting

De *toegepaste micro-econometrie* is het deelgebied van de economische wetenschap dat zich bezighoudt met het bestuderen van het economisch gedrag van individuen of huishoudens op basis van enquêtes. Dit proefschrift is een exponent van dit deelgebied. Het begrijpen van verschillende elementen van menselijk gedrag is belangrijk. Het geeft inzicht in hoe mensen kunnen reageren op veranderingen in politiek beleid, maar het kan ook helpen verklaren waarom sommige mensen rijk zijn en anderen arm. In dit proefschrift schat ik "equivalentieschalen", toets ik de validiteit van het levenscyclusmodel met behulp van subjectieve data, beschrijf ik patronen in belangrijke economische en demografische variabelen op basis van een representatieve enquête onder Nederlandse huishoudens, en bestudeer ik wanneer en hoeveel Nederlandse huishoudens werken, sparen en consumeren. Een belangrijke theoretische bijdrage van dit proefschrift is een aanzet tot het integreren van data over wat mensen doen –zogenaamde "revealed preferences data"– en wat mensen zeggen –zogenaamde "stated preferences data"– in één coherent framework.[1]

In hoofdstuk 2 schat ik equivalentieschalen op basis van een enquête onder Indonesische huishoudens.[2] Equivalentieschalen zijn huishoudenspecifieke wegingsfactoren die het inkomen, of het consumptieniveau van huishoudens *schalen* tot een consistentie welvaartsmaat. Het idee is dat een huishouden van twee personen meer inkomen nodig heeft dan een huishouden van één persoon om hetzelfde welvaartsniveau te bereiken, maar in het algemeen niet twee keer zo veel. Beide huishoudens

---

[1] Waar in deze samenvatting over "ik" wordt gesproken moet in acht worden genomen dat drie van de vier inhoudelijke hoofdstukken van dit proefschrift tot stand zijn gekomen in samenwerking met coauteurs: hoofdstuk 2 met Rob Alessie en Menno Pradhan, en hoofdstuk 4 en 5 met Rob Alessie.

[2] Ik maak gebruik van data van Indonesias National Socio-Economic Survey (SUSENAS)

hebben bijvoorbeeld maar één auto nodig en één huis. Desalniettemin kun je een brood maar één keer opeten. Een huishouden van twee heeft misschien E keer zoveel inkomen nodig als een huishouden van één, om hetzelfde welvaartsniveau te bereiken. De factor E is een equivalentieschaal en ligt waarschijnlijk ergens tussen 1 en 2. Door het inkomen van tweepersoonshuishoudens te schalen met de factor E (oftewel te delen door E) kan men het inkomensniveau *in welvaartstermen* vergelijken met het inkomensniveau van eenpersoonshuishoudens. Equivalentieschalen spelen een essentiële rol bij het berekenen van armoede statistieken (welk huishouden leeft boven en welk huishouden leeft onder de armoedegrens?) en bij het construeren van maten van welvaartsongelijkheid.

Het is niet onwaarschijnlijk dat equivalentieschalen afhankelijk zijn van de heersende prijzen en van het welvaartsniveau zelf. Binnen een parametrisch kader worden condities afgeleid waaronder equivalentieschalen onafhankelijk zijn van prijzen en/of het welvaartsniveau. Deze condities worden vervolgens statistisch getoetst en verworpen. De statistische verwerping van deze condities suggereert dat equivalentieschalen afhangen van prijzen en het welvaartsniveau. Echter, de omvang van deze afhankelijkheden zijn relatief klein en voor praktische toepassingen wellicht van ondergeschikt belang.

Met een eenpersoonshuishouden als referentiehuishouden worden de volgende equivalentieschalen geschat: 1.5 voor een tweepersoonshuishouden[3] (deze schatting is gelijk aan de veelgebruikte "OECD modified scale"), 2.0 voor een driepersoonshuishouden (deze schatting ligt duidelijk hoger dan de "OECD modified scale" die 1.8 rapporteert), 2.4 voor een vierpersoonshuishouden (de "OECD modified scale" rapporteert 2.1) en 2.7 voor een vijfpersoonshuishouden (de "OECD modified scale" rapporteert 2.4). Grosso modo zijn de equivalentieschalen iets groter voor grotere huishoudens dan de "OECD modified scales". Met andere woorden, de "OECD modified scales" *overschatten* het welvaartsniveau van grotere huishoudens (onder de aanname dat onze schattingen correct zijn).

In hoofdstuk 3 maak ik gebruik van subjectieve data ("stated preferences data") om de validiteit van het levenscyclusmodel statistisch te toetsen. Het levenscyclusmodel is een belangrijk theoretisch raamwerk dat in essentie tot doel heeft het spaargedrag van individuen te verklaren. Het model gaat ervan uit dat individuen rekening houden met toekomstige behoeften wanneer ze vandaag beslissingen nemen. Een eenvoudige versie van het model voorspelt dat men *spaart* wanneer het inkomen hoog is, en *leent* (of *ontspaart*) wanneer het inkomen laag is om consumptie optimaal te spreiden over de levensloop. In werkelijkheid echter, zijn consumptie (uitgaven) en inkomen beide relatief laag aan het begin van de levensloop en beide relatief hoog

---

[3] Dus, een huishouden van twee personen heeft 1.5 keer zoveel inkomen nodig als een huishouden van één persoon, om hetzelfde welvaartsniveau te bereiken.

rond middelbare leeftijd.[4] Deze en andere schijnbare tegenstrijdigheden vormen de basis van statistische verwerpingen van het model. Het levenscyclusmodel wordt echter in de literatuur vaak "gered" door toe te staan voor demografische effecten en "gewoontevorming": kinderen zijn duur en preferenties passen zich aan nieuwe omstandigheden aan.

Dit hoofdstuk introduceert een nieuwe fundamentele manier van het statistisch toetsen van het levenscyclusmodel (én economische modellen in het algemeen) dat ongevoelig is voor dergelijke generalisaties van het basismodel. De essentie van de methode is dat bepaalde elementen van de theorie direct worden gemeten met behulp van subjectieve data[5], zonder tussenkomst van willekeurige, en dus grillige, aannames omtrent functionele vorm. (Bijvoorbeeld: het introduceren van demografische effecten en gewoontevorming zijn uitbreidingen van functionele vorm.)

Het levenscyclusmodel kan worden getoetst door het toetsen van de validiteit van een eerste orde voorwaarde van het maximeringsprobleem, de Euler vergelijking van consumptie.[6] In de literatuur worden in het algemeen, op basis van functionele vorm aannames omtrent de *marginale* nutsfunctie, simultaan de toetsstatistieken en de modelparameters geschat. In het kort heeft deze empirische strategie twee nadelen. 1. De procedure test *slechts* de validiteit van een unieke parametrisatie van het model, niet het levenscyclusmodel in algemene zin.[7] Een levenscyclusmodel *in het algemeen* is een model waarbij de marginale nutsfunctie *niet* wordt gespecificeerd. 2. Door het simultaan schatten van modelparameters en teststatistieken is het niet eenvoudig om onderscheid te maken tussen algemene varianten van het levenscyclusmodel (een levenscyclusmodel met veel vrije parameters), en alternatieve theorieën over gedrag.

Door de marginale substitutievoet van consumptie in twee opeenvolgende perioden direct te meten met behulp van subjectieve data [de geïnteresseerde lezer wordt verwezen naar hoofdstuk 3 voor een uitgebreide uitleg over de procedure] worden de twee zojuist beschreven problemen opgelost. De uiteindelijke bevindingen zijn wisselend: het levenscyclusmodel met constante discontovoeten wordt verworpen. Het levenscyclusmodel met individuspecifieke discontovoeten wordt slechts marginaal verworpen. Zoals gezegd kan de verwerping van het levenscyclusmodel niet worden toegeschreven aan ontoereikende functionele vorm aannames. De resultaten zijn

---

[4] Deze observatie is inconsistent met een eenvoudige versie van het levenscyclusmodel, omdat: 1. de stijging in het inkomen ten dele kan worden geanticipeerd en 2. huishoudens worden verondersteld de geanticipeerde groei in het inkomen te spreiden over meerdere perioden.

[5] Het individu wordt direct naar zijn of haar mening gevraagd.

[6] Het levenscyclusmodel is een (nuts)maximeringsprobleem van een huishouden, waarbij het huishouden wordt geacht het verwachte nut over de levensloop te maximeren, gegeven schaarste in inkomen.

[7] In essentie is deze toets een toets op de gezamenlijke validiteit van de principes van het levenscyclusmodel én de specifieke aannames omtrent functionele vorm. Het kan dus zijn dat het levenscyclusmodel correct is, maar de functionele vorm niet toereikend.

dus een duidelijke verbetering, dan wel aanvulling op standaard testprocedures. De oorzaak van de verwerping wordt in dit hoofdstuk niet verder onderzocht. Daarentegen is de in dit hoofdstuk beschreven procedure is een toevoeging aan de literatuur met veel interessante en nieuwe openingen voor verder onderzoek.

Hoofdstuk 4 documenteert trends in belangrijke economische en demografische variabelen op basis van een representatieve[8] enquête onder Nederlandse huishoudens [bron: CBS Budgetonderzoek 1978-2000]. De variabelen die worden beschreven zijn duurzame en niet-duurzame consumptie, inkomen van het huishouden, werken of niet werken (man en vrouw), het aantal volwassenen per huishouden en het aantal kinderen per huishouden. De data wordt, met behulp van regressieanalyse, samengevat in leeftijds–, cohort (of generatie)– en tijdseffecten.[9] De leeftijds– en cohortprofielen worden vervolgens grafisch weergegeven.

De ontwikkelingen van deze variabelen over de levensloop (de leeftijdsprofielen) en over generaties (de cohortprofielen) zijn interessant om een aantal redenen. Ten eerste bieden ze uitgangspunten voor politieke en wetenschappelijke discussie. Voor interpretatie van deze profielen is een minimum aan theoretische kennis vereist, met als voordeel dat ze eenvoudig kunnen worden geïnterpreteerd en gecommuniceerd. Ten tweede bieden de profielen een ijkpunt voor macro-economische voorspelmodellen zoals het GAMMA model van het Centraal Planbureau dat onder andere wordt gebruikt om de verkiezingsprogramma's “door te rekenen”. Momenteel gebruikt het Centraal Planbureau de gerapporteerde leeftijdsprofielen uit dit hoofdstuk om de parameters van het GAMMA model te fixeren.

Tot slot bieden de leeftijdsprofielen op informele wijze inzicht in de validiteit van het levenscyclusmodel. Onder andere op basis van Amerikaanse data is een simultane stijging (in de eerste fase van de levensloop) en daling (vanaf middelbare leeftijd) van inkomen en consumptie gedocumenteerd, zélfs nadat wordt gecorrigeerd voor veranderingen in gezinssamenstelling. Dit heeft binnen de economische wetenschap tot veel discussie geleid, omdat het standaardmodel voor het analyseren van consumptie over de levensloop –het levenscyclusmodel– dit fenomeen moeilijk kan verklaren. Het is de vraag of deze patronen ook in Nederland belangrijk zijn. Verschillen in het gezondheidszorgsysteem en in de pensioenvoorzieningen tussen Nederland en de Verenigde Staten kunnen voor verschillende profielen zorgen.

Dit hoofdstuk documenteert onder andere dat inkomen, niet-duurzame en duur-

---

[8] De enquête is representatief voor praktische doeleinden. In sommige jaren zijn bepaalde types huishoudens oververtegenwoordigd. Voor deze overrepresentatie wordt gecorrigeerd doormiddel van weging.

[9] Een statistisch identificatieprobleem bemoeilijkt het maken van onderscheid tussen deze drie effecten. Door het opleggen van een parametrische restrictie op de tijdseffecten (deze “truc” is geïntroduceerd door Deaton and Paxson (1994)) wordt dit haalbaar. In de praktijk betekent dit dat alle “groei” wordt toegeschreven aan cohort effecten. De tijdseffecten kunnen vervolgens worden geïnterpreteerd als “business cycle”– of conjunctuureffecten.

zame consumptie *per huishouden* een duidelijk “bergvormig” patroon laten zien over de levensloop. We observeren een stijging in de eerste fase van de levensloop, vanaf leeftijd 20, gevolgd door een daling vanaf leeftijd 45. Dit patroon komt min of meer overeen met de bevindingen in andere ontwikkelde landen en lijkt universeel. Het is van theoretisch belang of de “bergvormigheid” in consumptie over de levensloop redelijkerwijs kan worden verklaard door een soortgelijk patroon in gezinssamenstelling. Met andere woorden, zijn de consumptieprofielen min of meer vlak (niet meer “bergvormig”) na het controleren voor veranderingen in de demografische samenstelling van het huishouden? De demografische patronen zijn bekend: net zoals consumptie en inkomen, neemt het aantal kinderen per huishouden toe tot leeftijd 45 en vervolgens af. Verder, het gemiddeld aantal volwassenen binnen het huishouden is praktisch constant tot leeftijd 60 waarna de kans dat een van beide volwassen leden van het huishouden overlijdt in toenemende mate belangrijk wordt.[10] Door inkomen, duurzame en niet-duurzame consumptie uitgaven te corrigeren met behulp van equivalentieschalen kijken we naar inkomen en consumptie *per volwassen equivalent*.[11]

De correctie voor gezinssamenstelling heeft een vrij dramatisch effect op de profielen. Inkomen vertoont nog steeds het bergprofiel, maar de top ligt later, rond leeftijd 60: huishoudens verdienen méér rond hun 45^{e} levensjaar, maar het inkomen moet met meer mensen worden gedeeld. De leeftijdsprofielen van duurzame en niet-duurzame consumptie daarentegen, laten beide geen duidelijke bergvorm meer zien. De verschillen tussen duurzame en niet-duurzame consumptie per volwassen equivalent zijn echter opvallend. Niet-duurzame consumptie per volwassen equivalent stijgt tot leeftijd 60 maar blijft daarna vrijwel vlak. Voor duurzame consumptie per volwassen equivalent zien we het omgekeerde: duurzame consumptie is praktisch vlak tot leeftijd 60 waarna het scherp daalt. In tegenstelling tot de resultaten uit Amerika, verdwijnt de bergvorm in beide maten van consumptie volledig na het controleren voor gezinssamenstelling. Echter, de simultane stijging van inkomen en niet-duurzame consumptie per volwassen equivalent blijft moeilijk te rijmen met de voorspellingen van een standaard levenscyclusmodel.

In hoofdstuk 5 modelleer ik het beslissingsprobleem om te gaan werken, te consumeren en te sparen. In dit onderzoek maak ik gebruik van dezelfde data als in hoofdstuk 4, echter, de empirische strategie is volledig anders. Waar de resultaten van hoofdstuk 4 zonder theoretisch kader kunnen worden geïnterpreteerd, zijn in dit hoofdstuk de resultaten volledig afhankelijk van het theoretisch kader. Dit maakt de interpretatie van de resultaten in zekere zin gemakkelijker. Door de data te bekijken

---

[10] De cohortprofielen laten een sterke toename in het aantal eenpersoonshuishoudens en een afname van het aantal kinderen per huishouden over generaties zien.

[11] Hiervoor gebruiken we de “OECD modified” equivalentieschaal. Equivalentieschalen nemen mee dat kinderen minder “verbruiken” dan volwassenen en dat er schaalvoordelen binnen het gezin te behalen zijn.

met behulp van een theoretisch fundament krijgen empirische patronen een duidelijke betekenis. Echter, eventuele fouten in het theoretisch raamwerk compromitteren dit voordeel. Het belangrijkste doel van het onderzoek is het schatten van arbeidsaanbodelasticiteiten voor vrouwen waarbij we rekening houden met de mogelijkheid om te sparen en te lenen.[12]

De economische theorie voorspelt dat de mogelijkheid om te lenen en te sparen een invloed heeft op de arbeidsaanbodelasticiteiten. Het typisch economisch mechanisme werkt als volgt: een stijging van het huidige uurloon maakt vrije tijd relatief duur ten opzichte van consumptie. Dit motiveert vrouwen om minder vrije tijd te consumeren en dus meer te gaan werken (om zo meer te kunnen consumeren). Als het een tijdelijke stijging van het loon betreft wordt vrije tijd *vandaag* ook duurder ten opzichte van de vrije tijd van *morgen*. Dit levert een extra motivatie om *vandaag* meer te werken. De extra opbrengst uit arbeid vandaag wordt verdeelt tussen vandaag en morgen door te sparen. Wanneer we rekening houden met de mogelijkheid tot sparen en lenen, schatten we de arbeidsaanbodelasticiteit van vrouwen op 1.7. In een statische omgeving, waarin we geen rekening houden met de mogelijkheid tot sparen en lenen, vinden we elasticiteiten van ongeveer 1.1.

Naast deze empirische bevindingen presenteert dit hoofdstuk ook een theoretische bijdrage aan de literatuur. Het standaard arbeidsaanbodmodel neemt aan dat huishoudens of individuen een efficiënte keuze maken tussen consumptie en vrije tijd. Een resultaat van dit keuzeproces is het aantal uren geleverde arbeid: het totaal aantal uren in een dag, vermindert met het aantal uren vrije tijd dat men kiest. Een belangrijk argument dat in dit model wordt weggelaten is de zorg voor kinderen of andere activiteiten binnen het huishouden. Dit noemt men in de economische wetenschap “domestic or home production” [vrij vertaald: werkzaamheden thuis]. Met andere woorden, individuen maken niet alleen een keuze tussen vrije tijd en consumptie, maar kiezen tussen vrije tijd, consumptie én de zorg voor kinderen thuis. Echter, dit hoofdstuk laat zien dat onder specifieke, maar redelijke functionele vorm aannames omtrent de “domestic production function” standaard arbeidsaanbodmodellen nog steeds logisch kunnen worden geïnterpreteerd.[13] De desbetreffende aannames zijn: de uren “arbeid thuis” van de twee volwassenen in het huishouden zijn de enige productiefactoren, de productiefunctie wordt gekenmerkt door constante schaalopbrengsten, en de twee productiefactoren zijn perfecte substituten.[14]

De eindconclusie van dit proefschrift is tweeledig. De belangrijkste bijdrage is dat

---

[12] Arbeidsaanbodelasticiteit: hoeveel procent meer (of eventueel minder) ga je werken als je uurloon met één procent stijgt.

[13] Een “domestic production function” vertaalt inputs (in dit voorbeeld het aantal uren arbeid thuis, van de man en de vrouw) in outputs (onder andere kinderopvang).

[14] Dit betekent overigens niet dat beide volwassenen even “productief” zijn. Het betekent slechts dat de marginale substitutievoet tussen de twee productiefactoren constant is.

“satisfaction data” (“stated preferences data”) een belangrijke rol kan, en zou moeten spelen in de economische wetenschap. In algemene zin kan subjectieve data als *extra* informatiebron worden gebruikt om algemeen aanvaarde, maar waarschijnlijk incorrecte vooronderstellingen in econometrische analyses van gedrag, te versoepelen of te toetsen.[15] Het toekomstperspectief is een completer beeld van individuele preferenties en een beter inzicht in gedrag. Het potentieel van “satisfaction data” (of andere soorten subjectieve data) in econometrische analyses hangt echter in belangrijke mate af van een juiste interpretatie van deze data. De eerste stap in het vervolgonderzoek is het doorontwikkelen van de in dit proefschrift beschreven theorie die “satisfaction data” interpreteert als een maat voor *marginaal nut* in tegenstelling tot een maat voor *nut*.

Verder concludeer ik in dit proefschrift dat het schatten van zogenaamde structurele economische modellen niet altijd tot betrouwbare inzichten leidt. Deze modellen zijn bruikbaar en geven richting aan de interpretatie van bepaalde patronen in de data. Echter, wanneer er geen wetenschappelijke consensus over de validiteit van zulke structurele modellen kan worden bereikt, verliezen de (geschatte) parameters van deze modellen hun originele interpretatie. Het is in deze gevallen aan te raden om minder krampachtig vast te houden aan het theoretisch kader en de theorie slechts te gebruiken als leidraad. Het rapporteren van geschatte structurele parameters is dan ook vaak minder informatief dan het rapporteren van eigenschappen van de data die, ook als de theorie onvolledig is, eenvoudig kunnen worden geïnterpreteerd. De hoofdstukken 4 en 5 benadrukken het verschil tussen de respectievelijk strikte en de lossere empirische benadering.

---

[15] Een aantal van die onwaarschijnlijke vooronderstellingen zijn rationele verwachtingen, de constantheid van preferenties over de tijd, de constantheid van preferenties tussen individuen. Deze lijst kan overigens eenvoudig worden uitgebreid.

Bibliography

Ahmed, N., Brzozowski, M. and Crossley, T. F.: 2007, Measurement errors in recall food consumption data, *http://www.econ.cam.ac.uk/faculty/crossley/research/abc_aug07.pdf* .

Alan, S., Attanasio, O. and Browning, M.: 2009, Estimating euler equations with noisy data: two exact gmm estimators, *Journal of Applied Econometrics* **24**(2), 309–324.

Alan, S. and Browning, M.: 2003, Estimating intertemporal allocation parameters using simulated residual estimation, *CAM working paper* .

Alessie, R. and De Ree, J.: 2009, Explaining the *hump* in life cycle consumption profiles, *De Economist* **157**(1), 107–120.

Altug, S. and Miller, R. A.: 1990, Households choices in equilibrium, *Econometrica* **58**(3), 543–70.

Amemiya, T.: 1985, Instrumental variable estimator for the nonlinear errors-in-variables model, *Journal of Econometrics* **28**(3), 273–289.

Apps, P. and Rees, R.: 1997, Collective labor supply and home production, *Journal of Political Economy* **105**(1).

Apps, P. and Rees, R.: 2003, Life cycle time allocation and saving in an imperfect capital market, *IZA Discussion paper. no. 1036* .

Attanasio, O. P.: 1998, Cohort analysis of saving behavior by u.s. households, *The Journal of Human Resources* **33**(3), 575–607.

Attanasio, O. P. and Weber, G.: 1995, Is consumption growth consistent with intertemporal optimization? evidence from the consumer expenditure survey, *The Journal of Political Economy* **103**(6), 1121–1157.

Auerbach, A. and Kotlikoff, L.: 1987, *Dynamic Fiscal Policy*, Cambridge Univesity Press.

Banks, J., Blundell, R. and Lewbel, A.: 1997, Quadratic engel curves and consumer demand, *Review of Economics and Statistics* **79**(4).

Banks, J., Blundell, R. and Tanner, S.: 1998, Is there a retirement-savings puzzle?, *The American Economic Review* **88**(4), 769–788.

Bean, C. R.: 1986, The estimation of "surprise" models and the "surprise" consumption function, *Review of Economic Studies* **53**(4), 497–516.

Becker, G. S.: 1965, A theory of the allocation of time, *The Economic Journal* **75**(299), 495–517.

Bernheim, B. D., Skinner, J. and Weinberg, S.: 2001, What accounts for the variation in retirement wealth among u.s. households?, *The American Economic Review* **91**(4), 832.

Beshears, J., Choi, J. J., Laibson, D. and Madrian, B. C.: 2008, How are preferences revealed?, *Journal of Public Economics* **92**, 1787–1794.

Blanchflower, D. G. and Oswald, A. J.: 2004, well-being over time in britain and the usa, *Journal of Public Economics* **88**, 1359 1386.

Blundell, R., Browning, M. and Meghir, C.: 1994, Consumer demand and the life-cycle allocation of household expenditures, *Review of Economic Studies* **61**, 57–80.

Blundell, R. and Lewbel, A.: 1991, The information content of equivalence scales, *Journal of Econometrics* **50**, 49–68.

Blundell, R. and Walker, I.: 1986, A life-cycle consistent empirical model of family labour supply using cross-section data, *Review of Economics Studies* .

Browning, M. and Lusardi, A.: 1996, Household saving: micro theories and micro facts, *Journal of Economic Literature* **34**(4), 1797–1855.

Browning, M. and Meghir, C.: 1991, The effects of male and female labor supply on commodity demands, *Econometrica* **59**(4).

Campbell, J. Y.: 1987, Does saving anticipate declining labor income? an alternative test of the permanent income hypothesis, *Econometrica* **55**(6), 1249–73.

Carroll, C. D.: 2001, Death to the log-linearized consumption euler equation! (and very poor health to the second-order approximation), *Advances in Macroeconomics* **1**(6).

Carroll, C. and Summers, L. H.: 1989, Consumption growth parallels income growth: Some new evidence, *NBER working papers* (3090).

Chamberlain, G.: 1984, *Handbook of econometrics*, Vol. II, Elsevier, Amsterdam, chapter Panel data.

Chiappori, P.-A.: 1997, Introducing household production in collective models of labor supply, *Journal of Political Economy* **105**(1), 191–209.

De Ree, J. and Alessie, R.: 2008, Home production and the allocation of time and consumption over the life cycle, *Netspar discussion paper* (020).

De Ree, J., Alessie, R. and Pradhan, M.: 2009, Prices, perceptions and the estimation of equivalence scales for indonesia, *unpublished* .

De Vos, K. and Zaidi, M. A.: 1997, Equivalence scale sensitivity of poverty statistics for the member states of the european community, *Review of Income and Wealth* **43**(3).

Deaton, A.: 1990, Saving in developing countries: Theory and review, *World Bank Economic Review, Special Issue, Proceedings of the Annual World Bank Conference on Development Economics* .

Deaton, A.: 1991, Saving and liquidity constraints, *Econometrica* **59**(5), 1221–1248.

Deaton, A.: 1997, *The analysis of household surveys: a microeconometric approach to development theory*, The Johns Hopkins University Press.

Deaton, A.: 2005, Franco modigliani and the life cycle theory of consumption, *unpublished* .

Deaton, A. and Muellbauer, J.: 1980a, An almost ideal demand system, *American Economic Review* **70**(3), 312–326.

Deaton, A. and Muellbauer, J.: 1980b, *Economics and consumer behavior*, Cambridge university press.

Deaton, A. and Paxson, C.: 1994, Intertemporal choice and inequality, *The Journal of Political Economy* **102**(3), 437–467.

Donni, O.: 2008, Labor supply, home production and welfare comparisons, *Journal of Public Economics* **92**(7), 1720–1737.

Draper, D. and Amstrong, A.: 2007, Gamma, a simulation model for ageing, pensions and public finances. CPB document no. 147.

Dynan, K. E.: 2000, Habit formation in consumer preferences: Evidence from panel data, *The American Economic Review* **90**(3), 391–406.

Evers, M., de Mooij, R. and van Vuuren, D.: 2008, The wage elasticity of labour supply: a synthesis of empirical estimates, *De Economist* **156**(1), 25–43.

Fernández-Villaverde, J. and Krueger, D.: 2007, Consumption over the life cycle: facts from consumer expenditure survey data, *The Review of Economics and Statistics* **89**(3), 552–565.

Gossen, H. H.: 1854, translated in English in 1983, *The laws of human relations and the rules of human actions derived therefrom*, The MIT Press.

Hall, R. E.: 1978, Stochastic implications of the life cycle-permanent income hypothesis: Theory and evidence, *Journal of Political Economy* **86**(6).

Hayashi, F.: 1985, The permanent income hypothesis and consumption durability: analysis based on japanese panel data, *The Quarterly Journal of Economics* **100**(4), 1083–1113.

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

Kahneman, D. and Krueger, A. B.: 2006, Developments in the measurement of subjective well-being, *Journal of Economic Perspectives* **20**(1), 3–24.

Kahneman, D., Wakker, P. and Sarin, R.: 1997, Back to bentham? explorations of experienced utility, *The Quarterly Journal of Economics* **112**, 375–406.

Kalmijn, M. and Alessie, R.: 2008, Life course changes in income: An exploration of age and stage effects in a 15-year panel in the netherlands, *Technical report*.

Kalwij, A., Alessie, R. and Fontein, P.: 1998, Household commodity demand and demographics in the netherlands: a microeconometric analysis, *Journal of Population Economics* **11**, 551–577.

Kapteyn, A.: 1994, The measurement of household cost functions. revealed preference versus subjective measures, *Journal of Population Economics* **7**, 333–350.

Kapteyn, A., Alessie, R. and Lusardi, A.: 2005, Explaining the wealth holdings of different cohorts: Productivity growth and social security, *European Economic Review* **49**(5), 1361–1391.

Kapteyn, A., Kleinjans, K. and Van Soest, A.: 2009, Intertemporal consumption with directly measured welfare functions and subjective expectations, *Journal of Economic Behavior & Organization* **72**(1), 425–437.

Kapteyn, A. and Van Praag, B. M. S.: 1976, A new approach to the construction of family equivalence scales, *European Economic Review* **7**, 313–335.

Koulovatianos, C., Schöder, C. and Schmidt, U.: 2005, On the income dependence of equivalence scales, *Journal of Public Economics* .

Lewbel, A.: 1989, Household equivalence scales and welfare comparisons, *Journal of Public Economics* .

Manski, C.: 2004, Measuring expectations, *Econometrica* **5**.

Modigliani, F. and Brumberg, R.: 1954, *Post-Keynesian Economics*, New Brunswick, NJ. Rutgers University Press., chapter Utility analysis and the consumption function: an interpretation of cross-section data, pp. 388–436.

Moffitt, R.: 1993, Identification and estimation of dynamic models with a time series of repeated cross sections, *Journal of Econometrics* **59**, 99–123.

Neary, J. P. and Roberts, K. W. S.: 1980, The theory of household behaviour under rationing, *European Economic Review* **13**, 25–42.

O'Donoghue, T. and Rabin, M.: 1999, Doing it now or later, *The American Economic Review* **89**(1), 103–124.

Olken, B. A.: 2005, Revealed community equivalence scales, *Journal of Public Economics* **89**, 545–566.

Pollak, R. A. and Wales, T. J.: 1979, Welfare comparisons and equivalence scales, *American Economic Review* **69**(2), 216–221.

Pradhan, M. and Ravallion, M.: 2000, Measuring poverty using qualitative perceptions of consumption adequacy, *The Review of Economics and Statistics* **82**(3), 462–471.

Ravallion, M. and Bidani, B.: 1994, How robust is a poverty profile?, *World bank economic review* **8**(1).

Ray, R.: 1983, Measuring the cost of children, *Journal of Public Economics* **22**, 89–102.

Samuelson, P. A.: 1938, A note on the pure theory of consumer's behaviour, *Economica* **5**(17), 61–71.

Schwarze, J.: 2003, Using panel data on income satisfaction to estimate the equivalence scale elasticity, *Review of Income and Wealth* **49**(3).

Smith, R. J. and Blundell, R. W.: 1986, An exogeneity test for a simultaneous equation tobit model with an application to labor supply, *Econometrica* **54**(3), 679–686.

Sullivan, H. S.: 1953, *The interpersonal theory of psychiatry*, W. W. Norton & Co., New York, NY.

Theeuwes, J. and Woittiez, I.: 1992, Advising the minister on the elasticity of labour supply, *Research Memorandum, Leiden University* **92**(06).

Townsend, R. M.: 1994, Risk and insurance in village india, *Econometrica* **62**(3), 539–91.

Van Praag, B. M. S.: 1968, *Individual welfare functions and consumer behavior*, North-Holland, Amsterdam.

Van Praag, B. M. S.: 1991, Ordinal and cardinal utility: An integration of the two dimensions of the welfare concept, *Journal of Econometrics* **50**, 69–89.

Van Praag, B. M. S. and Van der Sar, N.: 1988, Household cost functions and equivalence scales, *Journal of Human Resources* **23**(2).

Verbeek, M. and Vella, F.: 2004, Estimating dynamic models from repeated cross-sections, *forthcoming in: Journal of Econometrics* .

Vissing-Jörgensen, A.: 2002, Limited asset market participation and the elasticity of intertemporal substitution, *Journal of Political Economy* .

Zeldes, S. P.: 1989, Consumption and liquidity constraints: An empirical investigation, *Journal of Political Economy* **97**(2), 305–346.

Ziliak, J. P. and Kniesner, T. J.: 2005, The effect of income taxation on consumption on labor supply, *Journal of labor economics* **23**(4).