Essays on Household Decisions

Maria Gustafsson

Thesis submitted for assessment with a view to obtaining the degree of Doctor of Economics of the European University Institute

Florence, July 2014

European University Institute
Department of Economics

Essays on Household Decisions

Maria Gustafsson

Thesis submitted for assessment with a view to obtaining the degree of Doctor of Economics of the European University Institute

Examining Board

Prof. Jérôme Adda, EUI, Supervisor
Prof. Arpad Abraham, EUI
Prof. Frederic Vermeulen, University of Leuven
Prof. Thomas Crossley, University of Essex

Abstract

This thesis considers the fact that the majority of households consists of two adults whose characteristics and preferences matter for the households' decisions.

The first chapter studies how an increase in the generosity of maternity leave payments affects parental labor supply, early child development, and the relative well-being of the parents considering that parents may have different preferences over outcomes and that the policy change may affect the parental bargaining positions. I develop and estimate a static cooperative Nash bargaining model of parental decision-making in the first period of the child's life and use the model to investigate how the decision-making changes with an increase in the leave payments. The results indicate that mothers will spend more time at home rather than in the labor market when the leave payments increase, but that the average early child development is not much affected. Furthermore, the policy shifts the bargaining positions within the household in favor of the father and, although both parents are better off from the policy change, the mother would be better off relative to the father without the increase in maternity leave payments.

In the second chapter we look closer at how the insurance value of marriage, represented by the correlation of shocks to individual incomes, varies over different groups in the population. We find that this value may be lower for more recent cohorts, and decrease with age and with higher education.

The third chapter builds on the second. We investigate the importance of intra-household risk-sharing through labor supply by testing the following prediction: A higher correlation of income shocks within the household implies a lower ability to insure income through spousal labor supply and should, all else equal, lead to higher asset accumulation of the household. Our results indicate that this prediction holds empirically, suggesting that households perceive spousal labor supply as an important income insurance.

Acknowledgements

During the years that I have been working on this thesis I have met many great people who have contributed a great deal to these pages coming together.

I would first of all like to thank my supervisors, Jerome Adda and Russell Cooper for encouraging me to go my own way in research even though I many times found it difficult. They have also contributed with much valuable support and advice during this whole time. I would also like to thank Luigi Guiso and Paolo Sodini for introducing me to the exciting world of household finance, as well as the researchers at the Riksbank for the two interesting months I spent there and for all the helpful input. Andreas Fagereng and Kjersti Naess Torstensen are two persons who have had a large impact on the thesis. Without the excellent help of Kjersti in writing code I might have spent another year on this, and with the great minds and patience of them both, very abstract ideas have been turned into something more concrete.

The thesis has also benefited from comments and suggestions from participants in the Microeconometrics working group at the EUI as well as from classmates and other colleagues who have always been willing to discuss both problems and ideas. Moreover, one cannot put enough emphasis on the importance of the assistants, Jessica Spataro, Marcia Gastaldo, Julia Valerio, and Lucia Vigna who, by keeping track on professors and myself, have helped push the development of the thesis forward.

Finally, while writing this thesis the support from good friends and family has been extremely important. Thank you for always being there for me, especially in the moments when everything (including myself) felt totally impossible. And to Carmine: Vedo semple un Gatto.

Introduction

This thesis has its basis in more recent work, foremost in family economics, which emphasizes that two-person households should not be analyzed as consisting of one individual. An important reason is that interactions of household members that may have important implications for policy are thereby disregarded. This thesis applies a two-person framework to decisions regarding maternity leave as well as precautionary savings, and shows that including the other individual in the analysis of household decisions has implications for both the relative well-being of parents in response to changes in maternity leave payments, as well as for the ability of the household to insure labor income.

The first chapter studies how an increase in the generosity of maternity leave payments affects parental labor supply, early child development, and the relative well-being of the parents considering that parents may have different preferences regarding the outcomes and that the policy change may affect the parental bargaining positions. I develop and estimate a static cooperative Nash bargaining model of parental decision-making in the first period of the child's life, where the threat points are represented by a non-cooperative state of the relation. Data on labor supply, income and child development for the first period of a child's life are obtained from the first wave of the Millennium Cohort Study, UK. The estimated model is used to investigate how the decision-making changes with an increase in the leave payments. The results indicate that mothers will spend more time at home rather than in the labor market when the leave payments increase, but that the average early child development is not much affected. Furthermore, the policy shifts the bargaining positions within the household. When the mother has relatively low preferences for working, her bargaining power increases relative to the father, whereas her relative bargaining power decreases when she has relatively high preferences for work. When the father gains in bargaining power the parents redistribute resources from child development to consumption in accordance with his preferences. Both parents are better off from the policy change, but overall the mother is better off relative to the father without the increase in maternity leave payments.

In the second chapter we look closer at, what may be called, the insurance value of marriage. In particular, we study the correlation of shocks to the income of individuals within a household. The higher this correlation is, the lower is the ability of the household to insure income through spousal income and labor supply. In the Chapter we are interested in whether this ability varies over groups in the population, as well as over time. We estimate

the correlation of income shocks using administrative data from Norway. This dataset is very well suited for these purposes as we are able to observe individuals and their incomes over long periods of time without much problems of measurement error. We find that the correlation of income shocks is higher for more recent cohorts and that it increases slightly over age, with a tendency of a hump-shape for some groups. Moreover, higher educated individuals tend to have higher correlation, although the variance in correlation seems greater for this group. This provides an indication that the nature of household income risk has changed over time and that it changes with age and with education levels. We also find that assortative matching on industry and education might be able to account for part of the correlation patterns.

Following up on these results, in the third chapter we investigate the potential importance of heterogeneity in income shock correlation in the population. More specifically, we test whether spousal labor supply is an important mechanism for the household when it comes to insure labor income. We do this by testing the prediction that if labor supply is important for insurance, the correlation in income shocks within the household should matter for precautionary savings, independent of the variance of the shocks. For the analysis we link the income information to information on individual asset holdings from the Norwegian tax registry. Our results indicate that the prediction holds empirically, suggesting that households in Norway use spousal labor supply to insure against labor income risk.

Contents

Chapter 1

Maternity Leave Payments and the Intra-household Distribution of Resources

1.1 Introduction

Early child development measures have been shown to be strong predictors of a child's educational outcomes.[1] Moreover, maternal time spent with the child in the first period of the child's life is considered important for how well the child will do later in life through mechanisms as for example breastfeeding, bonding with the child, and attention to the child's health.[2] As the labor force participation of women has increased a lot the last decades and they have become important breadwinners, maternity leave payments have become an important policy for many families. A payment in the leave period facilitates for the household to have the mother taking time off from work after birth to care for herself and the child.[3] As Figure 1.1 shows, all OECD countries except the U.S. have a public system

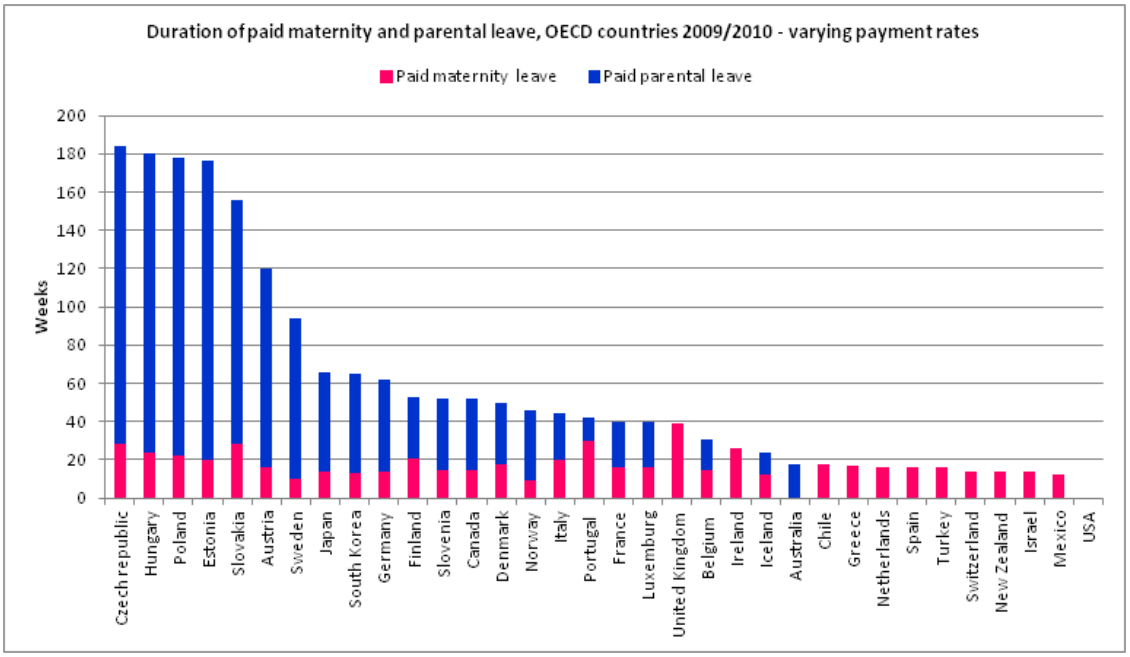
[1] Examples are Bernal and Keane (2008) who show that test scores at the ages of 4-6 are strongly correlated with the educational attainment; and the findings by Carneiro, Crawford and Goodman (2007) which indicate that early indicators of development such as whether the child walked at age 1.5, talked by age 2, and still wet him/herself at age 3, are strong predictors for the child's cognitive and social abilities at age 7. Aizer and Cunha (2012) show that Bayley test scores at 8 months can predict achievements in math at age 7.

[2] See for example Ruhm (2000) and references therein.

[3] The focus of this paper is on payments directed to the mother. Most countries also have *parental* leave payments, which are payments that can be taken by both parents. However, mothers are still the ones taking the vast majority of the parental leave.

of paid parental leave in place.[4] Even though the duration of paid leave varies a lot among the countries the majority has around six months of leave payments.

Figure 1.1: Duration of paid leave in the OECD[5]



From Figure 1.2 we can see that the duration of leave payments has increased over time. A common motivation behind increases in the payment generosity is the importance of mother time and care for the welfare of the child and that the payments increase the welfare of the family, and in particular that of the mother.[6]

However, it is not clear from current research that a more generous leave policy has positive effects on the child's future outcomes (e.g Baker and Milligan, 2008, 2010, 2011;

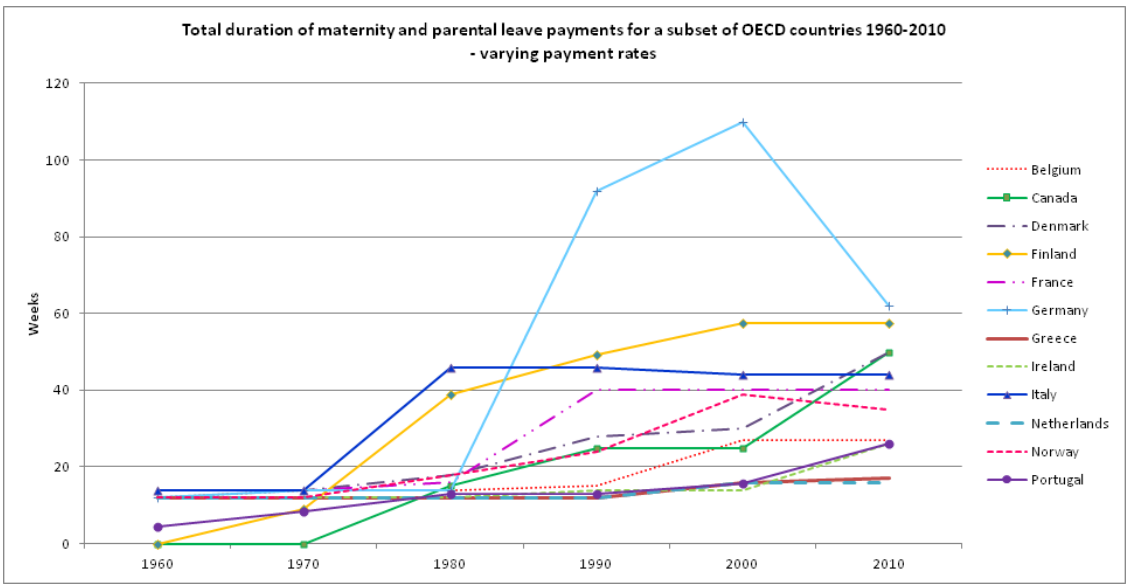
[4]Although California, New Jersey and Rhode Island have implemented paid family leave insurance programs.

[5]The information used in the figure is from the appendix of report "Failing its Families: Lack of Paid Leave and Work Family Supports in the US" from the Human Rights Watch. Available at http://www.hrw.org/node/96432. When the information was given in days I have divided by five and multiplied by four. When the information was given in years I have multiplied by 12. In the case when an option of a longer leave duration but for lower pay was possible I have used the short period with the high payment.

[6]Dustman and Schnberg (2012) and Liu and Nordstrm Skans (2010) note that the welfare of the children has been an important argument behind increases in the generosity of parental leave policies in Germany and Sweden respectively.

[7]Most of the information used in the figure is from Gauthier (2011) "The Comparative Family Policy Database" Version 3. In the case of Germany, also from Dustmann and Schnberg (2012). Other sources of information: Ruhm and Teague (1995); The Clearinghouse on International Developments in Child Youth and Family Policies, at Columbia University.

Figure 1.2: Duration of paid leave over time[7]



Carneiro, Løken, and Salvanes, 2011; Liu and Nordstrm Skans, 2010; Rasmussen, 2010; Dustmann and Schnberg, 2012; and Dahl, Løken, Mogstad and Salvanes, 2013). In general the researchers find that the labor supply of the mother is very responsive to changes in maternity leave policies, but that there is hardly any effects on short-and long-run child development (with the exception of Carneiro et al., 2009).[8]

Moreover, the effects of paid maternity leave policies on parental decision-making and intra-household resource distribution are not well understood. Consider that parents have different preferences, that decisions are taken by bargaining, and that a change in the maternity leave policy may affect the bargaining positions of the parents. In this case the change in maternity leave does not only affect the decisions through income and substitution effects, but also through a change in the household objective function where the preferences of the parent who gained in bargaining power will be weighted higher. Consequently there would also be an effect on the distribution of well-being between the parents.

In this paper I seek to investigate these issues by estimating a one-period model of parental decisions regarding time and income allocations. The period in the model is the first nine months of a child's life and the estimated model is used to analyze the effect of a more generous maternity leave payment on the child development and on the intra-household

[8]There is an extensive literature investigating the effect of mother work early in the child's life on the child's later outcomes, see for example Brooks-Gunn, Han, and Waldfogel (2002), Berger, Hill and Waldfogel (2005), Ruhm (2008). A critical review of the studies using U.S. NLSY data can be found in Bernal and Keane (2010).

distribution of resources in this period. The main scarce resource in the model is time which the parents choose to spend in the labor market or at home. Time at home contributes to child development and time in the labor market generates income used for consumption and for buying non-parental childcare. Following the seminal work by Manser and Brown (1980) and McElroy and Horney (1981) the parents make decisions through cooperative Nash bargaining.[9] Importantly, this set-up allows for the parents to have different preferences over child development, consumption and their own working time. What the household will weight higher depends on the interaction of preferences and the threat points of the parents which I will refer to as their bargaining positions.[10] The threat points are determined by the utility the parents would attain if they got into a disagreement, modeled as a non-cooperative Nash equilibrium (e.g. Lundberg and Pollak, 1993; Chen and Wolley, 2001; Lechene and Preston, 2011).[11]

The child development observed at the end of the period is, apart from parental time, a function of initial child development, the household consumption level and of the time spent in non-parental childcare. To account for systematic differences of mothers who work more that are related to the child development, as well as for the possibility that parents with children of low initial ability may choose to spend more time at home, the child development function is estimated jointly with the model of household choices.[12]

The model is estimated using a Simulated Method of Moments approach. With this approach the parameters of the model are chosen such that moments generated by the model match moments from data. The data is obtained from the first wave of the U.K. Millennium Cohort Study which is a longitudinal survey of children born in the U.K. around the millennium. This survey is particularly good for the current study as it provides detailed

[9]The household decision-making framework that allows for different preferences and weights of the family members in the household utility function has been favored by a number of empirical tests (e.g Browning et. al., 1994; Lundberg et. al., 1997; Phipps and Burton, 1998; Lundberg et. al., 2003; Alessie et. al., 2006; Vermeulen, 2005; and Cherchye, De Rock, and Vermeulen, 2009.). For a reviews of the unitary and the collective approach to household modeling, see for example Vermeulen (2002) and Browning, Chiappori, and Lechene (2006).

[10]Two papers studying the effects of parental leave policies on the aggregate economy using general equilibrium models are Bernal and Fruttero (2008) and Erosa, Fuster, and Rusticcia (2010). Bernal and Fruttero (2008) use a bargaining framework for household decision-making, but do not explicitly consider changes in intra-household bargaining positions from maternity leave policies.

[11]The seminal papers by Manser and Brown (1980) and McElroy and Horney(1981) (as well as many other following them) assume that the threat points are determined by the utility the parents would obtain in case of divorce. Here it is not divorce that is the most credible threat, but a stage (potentially just before divorce) when the parents act non-cooperatively.

[12]Examples of studies estimating a production function of child welfare jointly with an explicit model of household choices are Bernal (2008), Del Boca, Flinn, and Wiswall (2011), and Cherchye, De Rock and Vermeulen (2012).

information on the characteristics of the parents, their labor supply in the very first period, and measures of child development when the child is around nine months.

The estimated model fits the data well. The parameter estimates show some interesting features and contribute to insights about household preferences and child development in the very first period of a child's life. For example, the estimations indicate that the mother puts a higher weight on child development than the father, and that fathers dislike work less than mothers in general. Among both fathers and mothers there are types with high and low preferences for working. Low preference parents dislike work, but the high type parents have positive preferences for work. Regarding child development, the estimations show that initial ability is very important for the development at nine months of age, and that the initial ability is strongly associated with unobserved variables of the parents. Moreover, an hour at home of the mother is estimated to be about twice as important for the child development as an hour at home of the father and the parents are found to be slightly complementary in child development production. As compared to parental time in the child development function, the importance of both the level of income and the amount of non-parental childcare is found to be fairly low.

The model is used to simulate the impact of a change in the maternity leave payment policy that was actually implemented in the UK in 2003. The change in policy featured an increase in the duration and in the payments of the maternity leave payment. I simulate the impact on the distribution of resources within the household and in particular I look at changes in parental labor supply, child development, and the relative well-being of the parents. The results suggest that the labor supply of both parents decreases as compared to the situation under the original policy, although the impact is much greater for the mothers (around 20 as compared to 4 percent). However, despite the decrease in labor supply and an average increase in income in the sample, there is only a small overall increase in child development (1.5 percent). Moreover, as parental time is estimated to be more important for child development than income, and low income mothers are less likely to work under the original policy, the average increase in child development is lower for the low income households than for the high income households.

Since I have estimated the parameters of the model I can simulate the behavior of the parents if they would not cooperate. When introducing the new policy I find that the bargaining power of the mothers seems to increase on average. However, when comparing the high preference type mothers to the low types I find that mothers who enjoy working more loose in bargaining power. This is largely due to a free-riding effect in the non-cooperative equilibrium. The mother would work less also when not cooperating which contributes to

the child development, a contribution the father can enjoy without decreasing labor supply and consumption. Similarly the low preferences mothers are happier increasing their time at home. Following up on this result I simulate the outcomes in the cooperative situation keeping the bargaining power fixed. The results suggest that the effect of the bargaining power change is important for the relative well-being of the parents and that the average results are driven by the high preference mothers. The increase in the bargaining position of the father pushes the household decisions towards more father work, more consumption and less child development, in line with the father's preferences. In the case where the bargaining power changes the mother is better off from the policy by 1470 pounds and the father by 960 pounds. In the case where the bargaining power is held fixed the same numbers are 1750 and 800. These numbers masks some differences among types. The low preference mothers are better off when there is a bargaining power change by around 300 pounds, whereas the high preference mothers are worse off when the bargaining power changes by around 600 pounds.

The conclusion from these results is that a fairly big increase in the generosity of leave payments allows for better consumption possibilities and facilitates for the mother to stay at home with the child after its birth. However, the policy change does not have much effect on the child development and it changes the bargaining positions of the parents within the household, where mothers who are stronger attached to the labor market loose in bargaining power relative to their partner and vice versa. Thus, both parents are better off from the increase in leave payments, but the mother with a stronger attachment to the labor market is worse off relative to the father.

The paper proceeds as follows. The model is presented in Section 1.2, and in Section 1.3 I describe institutional background and the data used for the estimation. Section 1.4 lays out the estimation strategy as well as the estimation results. Section 1.5 contains the policy analysis and Section 1.6 concludes.

1.2 A one period model of parental decision-making

1.2.1 Set-up

In the model there is a father (f), a mother (m) and a newborn child. In the following a parent will be denoted by p, where $p = m, f$. Both parents have time endowment T which is the maximum hours of work in this period. The parents derive utility from seeing their child develop and the development is summarized in the variable Q. They also care about consumption, c and hours of market work, m_p. Moreover, the parent might enjoy being at

home more if the partner is at home more and therefore there is an interaction term in the model, consisting of the home time of each of the parents. Q is a public good whereas there is a more complicated set-up for consumption explained below.

The preferences of the parents may differ and they are specified in the following way:

$$U_p = \frac{c^{\delta_c}}{\delta_c} + \alpha_p \frac{Q^{\delta_q}}{\delta_q} - \nu_p(a_p)m_p + \nu_{p-p}h_p h_{-p} \tag{1.1}$$

where $\delta_c \leq 1,\ \delta_q \leq 1$. Important to note is that the parents may have differing preferences regarding the relative importance of consumption, child quality, and work hours. Furthermore, there is unobserved heterogeneity in preferences for working. The preference type is indicated by a_p and will be discussed below.

The parents face a constraint set consisting of time and goods expenditure restrictions. The individual time endowment, T, can be used for labor market work, m_p or for spending time at home, h_p. That is, $m_p + h_p = T$. The time spent in the labor market generates income $w_p m_p$, where w_p is the hourly wage rate of the parent. The mother may be eligible for income compensation during the leave time which is expressed by $B(M_m, w_m, week_m)$. The payments depend on the income in the period before birth, summarized by how much the mother worked in the period before birth, M_m, and her wage rate, w_m.[13] As maternity leave payments are paid on a weekly basis, also the number of weeks becomes a choice variable of the mother, denoted by $week_m$.

The income of the parents can thus be expressed as

$$I_m = w_m m_m + B(M_m, w_m, week_m) \tag{1.2}$$

$$I_f = w_f m_f. \tag{1.3}$$

Note that the only thing that is important for the father income is the total amount of hours, and not how they are distributed over weeks.

Non-parental childcare, oc, is modeled as a discrete choice. If the family chooses to use non-parental childcare, this covers the hours the mother is at work during this period. Otherwise it is assumed that the parents arrange their working time such that the father looks after the child. The cost of non-parental childcare is p_{oc} per hour which includes both the monetary and potentially non-monetary costs for the parents of not taking care of the child themselves. It should be noted that the model does not distinguish between different

[13]The assumption here is that what the mother can earn in the labor market has not changed drastically from just before the birth to just after.

types of childcare in order for the model to be more tractable. However, studies show that different types of childcare may be of different quality (e.g. Brilli, Del Boca, and Monfardini, 2013). If some mothers are constrained to use the highest quality childcare because of the price is too high, for example, this may create problems in the parameter estimation. This is further discussed in Section 1.3.2.

Finally, as there are no savings in the model, all income that is not spent on non-parental childcare will be used for consumption. The household budget constraint:

$$c + p_{oc}oc = w_f m_f + w_m m_m + B(M_m, w_m, week_m) \tag{1.4}$$

The parental time that is not spent in the labor market serves as the input into a child development production function. Further, non-parental childcare and the level of consumption also contribute to the child's development. The child development function is assumed to be:

$$Q = q_0(bw, e, a_f, a_m) + \left(\kappa_m(e)h_m^{\phi} + \kappa_f h_f^{\phi}\right)^{1/\phi} + \kappa_c \frac{c^{\phi_c}}{\phi_c} + \kappa_{oc}oc \tag{1.5}$$

where ϕ summarizes how well the mother and father input can be substituted in the production and ϕ_c denotes the decrease in the marginal productivity of consumption. $\phi \leq 1$, $\phi_c \leq 1$. The return to outside care is assumed to be linear. κ_f and κ_m denotes the importance of an hour not spent at work of each parent respectively, κ_c is the productivity of one unit of household consumption, and κ_{oc} is the productivity of one hour of non-parental care. In order to be consistent with the bargaining framework it would be desirable to subject also the consumption spent on the child to bargaining between the parents. However, as the data do not provide any information on expenditures on children (clothing, food, toys, etc.) this is not possible in this study.[14] Furthermore, one may think of the consumption level of the parents as indicating other things that may be more important for how well the child develops, as for example the environment the child grows up in, whether the family has heating, etc.

According to existing theories, mother education may be important for child development for at least two reasons. The first is through an effect on initial child development which may come from a better knowledge and possibilities of preparation before birth. For example when it comes to nutrition and the use of cigarettes, alcohol and coffee. Variables such as

[14] In their study, Cherchye, De Rock and Vermeulen (2012) are able to distinguish between expenditure on children, as well as private consumption of the parents.

prenatal care and maternal stress may also be correlated with education.[15] Secondly, studies have found indications that mothers with a higher education may be more productive in taking care of the child (e.g Liu and Nordstrom Skans, 2010). To allow the productivity of time spent with the child to be different for mothers with higher education I let κ_m vary with the education level of the mother, e.[16]

In the literature on child development there is much discussion around the endogeneity of inputs into the production function, especially in regards to hours of work and income.[17] Mothers that work more may be systematically different from other mothers in ways that are important to the child development. For example, if more able mothers would both work more and have more able children, an estimated effect of hours of work on child development would have a positive bias. If this "ability" is constant over time it could be accounted for using panel data and the fixed effects estimator. Moreover, there is at least one other concern that can not be taken care of using the fixed effects approach. If the development of the child at birth is very low, the mother may prefer to spend more time with the child rather than in the labor market. In this case time at home would be estimated to be less productive than it is. In order to overcome this bias in a reduced form model an instrument is needed. When using a structural model, however, it is possible to capture both types of endogeneity, also with cross-sectional data. The model in this study includes the inital development of the child q_0, i.e. the development at birth, which the parents (but not the econometrician) observe and based on this variable the parents make decisions regarding for example labor supply. The initial ability is estimated together with the other parameters and it is allowed to depend on the birth weight of the child, bw, the education level of the mother, e and the unobserved work preferences of the parents, a_m, a_f. The inclusion of the latter variables allows for an association between the time spent working of the parents and the development of the child.[18]

The initial child development is expressed as:

$$q_0 = \beta_0 + \beta_{bw} bw + \beta_e e + \beta_m a_m + \beta_f a_f. \tag{1.6}$$

[15]Currie and Moretti (2003) investigates the causal effect of education on important prenatal variables by using the introduction of a policy increasing the supply of colleges when the mother was a teenager. They find that education has a positive effect on the use of prenatal care and the birth weight, and a negative effect on smoking.

[16]This could also be true for the fathers, but evidence on the effect of father time spent with the child is as mentioned very scarce in general. I leave this out here in order to keep the state space smaller.

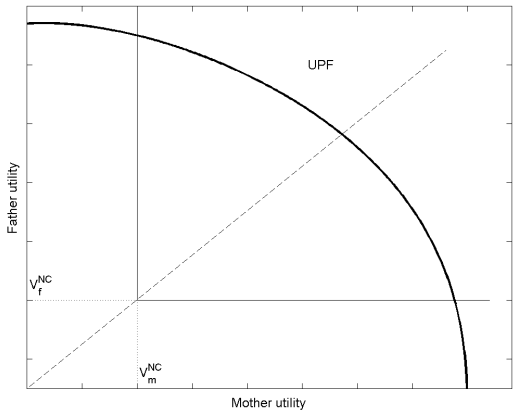
[17]See discussion in, for example Bernal (2008).

[18]In the estimation I will assume that this association is positive. This is done because an OLS regression using the variables in the data shows that hours of work of the parents are positively associated to the measure of child development.

1.2.2 Parental decision-making

The assumption on how the parents make decisions is an important feature of the model and deserves some additional explanation. The observed outcomes in terms of labor supply, child development, and consumption are assumed to be the result of cooperative decision-making, and more precisely cooperative Nash bargaining. Pareto optimality is achieved in that the solution lies on the utility possibility frontier of the household.[19] Exactly where on this frontier is determined by what would happen if the cooperation breaks down. If the cooperation breaks down the parents are assumed to solve the problem in a non-cooperative manner according to a non-cooperative Nash bargaining problem. The utility the mother and the father would obtain in the disagreement situation respectively is the individual's threat point in the cooperative problem. One can think of this as credible threats. If the mother would be better off than the father in case of no cooperation her threat of not cooperating is more credible and the father transfers some of his utility to her. For an illustration, see Figure 1.3. An important point is that the threat points are exogenous to the cooperative problem.

Figure 1.3: Utility possibility frontier when cooperating



There are two important differences between cooperative and non-cooperative parents. As mentioned above the preferences regarding consumption changes. When the parents are in a cooperative relationship the consumption is a public good. However, when the parents are in the non-cooperative relation they do not derive any utility from the other

[19]See for example Del Boca and Flinn (2012a,b) for discussion and empirical study of whether households can be said to behave efficiently.

parent's consumption and consumption takes on the characteristic of a private good.[20] That is, the parents more or less live two separate lives within the household. Moreover, in the cooperative case, there is a complementary effect of the other parents work on the disutility from working which implies that if the other parents work more, the first parents would want to work more him/herself. Altogether these assumptions have two important implications. First of all, the parents do not face the same utility possibility set in the two states and they would always prefer to cooperate. Secondly, that consumption is private when not cooperating ensures that wage is positively associated with bargaining power which seems plausible.[21]

In the following I will first describe the cooperative situation which I assume generates the outcomes we observe in the data. In this way I can present the importance of the threat points before I move on to describe how these are determined.

Cooperative parents

When cooperating the parents pool their incomes and jointly decide upon how much time that should be devoted to market work to generate income, how much time that should be devoted to home time to produce child development, and whether or not to use non-parental childcare when the mother is working. Consumption is a public good under cooperation and the budget constraint is the following:

$$c + p_{oc}oc = w_f m_f + w_m m_m + B(M_m, w_m, week_m) \tag{1.7}$$

According to the cooperative Nash solution concept, the choices of the parents are made in order to solve the following household problem:

$$\max_{c,Q,h_f,h_m,m_f,m_m,week_m} \tag{1.8}$$

$$\left(\frac{c^{\delta_c}}{\delta_c} + \alpha_f \frac{Q^{\delta_q}}{\delta_q} - \nu_f(a_f)m_f + \nu_{mf}h_m h_f - V_f^{NC}\right) \times \left(\frac{c^{\delta_c}}{\delta_c} + \alpha_m \frac{Q^{\delta_q}}{\delta_q} - \nu_m(a_m)m_m + \nu_{mf}h_m h_f - V_m^{NC}\right) \tag{1.9}$$

subject to the budget constraint, the production function and the time constraints. If the parents choose non-parental childcare the childcare hours will cover the hours the mother is working and the income would be used to cover consumption and the cost of non-parental

[20] The set-up of the non-cooperative game is closely related to the set-up in Konrad and Lommerud (1995).

[21] It would be interesting to model consumption as private also when cooperating. However, I cannot estimate the individual preferences for consumption as I do not have data on private consumption.

childcare. If not, the whole income will be used for consumption. As mentioned above the threat points, V_f^{NC} and V_m^{NC}, are exogenous to the problem.

The optimal parental working time is denoted by m_f^C and m_m^C for the father and mother respectively, the optimal weeks at home $week_m^C$, the choice to use non-parental childcare or not is d_{oc}^C, with the corresponding consumption and child development c^C and Q^C. The optimal working time, consumption and child development generates utility U_f^C and U_m^C from cooperating. The working time of the parents will depend on their preferences for consumption and child development and their productivity in each of the activities. If both parents are more productive in the labor market than in child development the household will tend to spend more time working in the market, all else equal. Similarly, if the parents have stronger preferences towards child development than consumption they will tend to spend more time in the home.

One of the more interesting feature of the model is that the bargaining power plays an important role in the time allocation. For example, if the parent with strongest preferences for child development also has a relatively strong bargaining power, the couple tend to spend more time with the child. If changes in the parental leave policy change the bargaining power of the parents it implies an indirect effect on the time allocation and the parent who's bargaining power increased relative to the other parent is relatively better off in the final allocation.

Non-cooperative parents

The threat points, or the bargaining power of the parents, depend on the optimal behavior in a disagreement situation which is assumed to be a non-cooperative Nash equilibrium. The assumptions regarding the non-cooperative environment are the following: The parent only consumes his or her own income; there are no transfers between the parents; and the mother is responsible for the non-parental childcare. The latter implies that if the mother would like to work she would also have to pay for the childcare herself. Thus, one can think of this environment as a stage just before a divorce.

The budget constraints are the following: $c_f = w_f m_f$ and $c_m + p_{oc} oc = w_m m_m + B(M_m, w_m, week_m)$ where the price of the consumption good is assumed to be the same for both parents and normalized to 1. Further, as mentioned above, the preferences are egoistic in the sense that one parent does not derive utility from the consumption of the other parent. In terms of decision-making the parents are assumed to play a simultaneous game of private contributions to one public good taking the other person's contribution as given.

The problem of the mother and the father looks slightly different as the mother makes the childcare choice and pays the costs. Moreover, the mother is also choosing over both hours and weeks.

The problem of a non-cooperating father

$$\max_{c_f,Q,h_f,m_f} \frac{c_f^{\delta_c}}{\delta_c} + \alpha_f \frac{Q^{\delta_q}}{\delta_q} - \nu_f(a_f)m_f \tag{1.10}$$

subject to

$$Q = q_0(bw, a, a_f, a_m) + \left(\kappa_m(e)\bar{h}_m^{\phi} + \kappa_f h_f^{\phi}\right)^{1/\phi} + \kappa_c \frac{((\bar{c}_m + c_f))^{\phi_c}}{\phi_c} + \kappa_{oc}\bar{oc} \tag{1.11}$$

$$c_f = w_f m_f \tag{1.12}$$

$$m_f + h_f = T. \tag{1.13}$$

The problem of a non-cooperating mother The problem of a non-cooperative mother is fairly similar to that of the father, although she also makes the choice of childcare or not. Thus:

$$\max_{c_m,Q,h_m,m_m,week_m} \frac{c_m^{\delta_c}}{\delta_c} + \alpha_m \frac{Q^{\delta_q}}{\delta_q} - \nu_{m,}(a_m)m_m \tag{1.14}$$

subject to her budget constraint, the production function and the time constraint. The choice of childcare implies that her income will cover childcare costs and consumption. Otherwise the mother consumes the whole income.

Solving for one parent's optimal labor supply in terms of the other parent's labor supply we get the best response function of this parent. In an interior solution their intersection determines the Nash equilibrium of the game. A Nash equilibrium in the model is a pair of labor supplies, m_f^{NC}, m_m^{NC}, a choice of weeks at home $week_m^{NC}$, and the choice of using non-parental childcare, d_{oc}^{NC}, such that these are mutually optimal responses and yield quantities c_f^{NC}, c_m^{NC}, and Q^{NC}. The optimal time allocation of the parents will depend on their preferences for consumption versus child development as well as their productivity in each of the activities. Since consumption is a private good and child development is a public good there is a tendency to free-ride on the partner's public good contribution. All else equal the parent with a relatively high productivity in the labor market will work more and be able to consume more while still enjoying the child development contribution of the parent who

is relatively productive in child development.[22]

The optimal working time, consumption and child development generates utility V_f^{NC} and V_m^{NC} which represent the threat points in the cooperative bargaining set-up presented below.

At this point we cannot make clear predictions about how an increase in the payment generosity would affect the utility in the non-cooperative equilibrium as not even the non-cooperative problem is possible to solve analytically without making restrictive assumptions regarding the parameter values. As the income is private, an increase in maternity leave payments would make the mother better off as her possibility set has expanded. However, her choices regarding time at home and childcare also benefits the father and there could be a free-riding effect in case the best response of the mother is to increase the contribution.

1.2.3 Model solution

As it is not possible to solve the model analytically the model will be solved numerically in the way described below.

A household with a newborn child is characterized by Ω_h. Ω_h includes the wages of each parents (w_f, w_m), the working time in the period before the child was born (M_m), which determines the amount of leave payments the mother is entitled to, the education level of the mother (e), the birth weight of the child (bw), and the unobserved heterogeneity of the parents (a_f, and a_m). That is: $\Omega_h = (w_f, w_m, M_m, e_m, bw, a_f, a_m)$.

In order to solve the model I discretize the characteristics and the outcome variables in the following way. In accordance with the data, which shows a larger spread of the mothers' wages, the wages belong to the sets $w_f \in \{3, 4, ..., 12\}$ and $w_m \in \{3.5, 4.5, ..., 12.5\}$.

The hours of work of the mother in the period before birth will first of all determine whether she is eligible for leave or not. Then, if the mother is entitled to leave payments the amount she receives depends on her weekly income in the period before the birth of the child (see Section 1.3.2). Assuming that the mother's wage is about the same in the period before the child is born as in the immediate period after, the approximate weekly income in the last period can be calculated using the wage and information on whether the mother worked half-time or full-time in the period before birth. Consequently the hours for which the mother can claim benefits, M_m, can be restricted to the set $\{0, 20, 40\}$ where 0 is when the mother is not eligible for any leave payments and 40 is full-time. More on how I calculate the leave payments in the model and in the data can be found in Section 1.3.2.

[22] In this case the public good is child development and the private contributions are most importantly time. For an example of free-riding in a similar family context see Konrad and Lommerud (1995).

The mother's productivity in childcare as well as the initial child development are allowed to vary with the mother's education which can be high or low.

The unobserved heterogeneity, a_p is assumed to take two values: either high or low. Thus we have four different types: mother high, mother low, father high, and father low. This further implies that we have four different parental couple combinations: low-low; high-low; low-high; and high-high.

The birth weight is discretized into three categories. The child can be born with a normal birth weight, with a lower than normal birth weight, or with a birth weight higher than normal. Thus we can express the child's initial development as:

$$q_0 = \beta_0 + \beta_{bw}^l bw^l + \beta_{bw}^h bw^h + \beta_e e + \beta_m a_m + \beta_f a_f \tag{1.15}$$

Note that the type combination of the parents determines the initial development of the child by a contribution parameter of the father and the mother respectively that are estimated within the model.

The heterogeneity parameters are estimated together with the relative parental type proportions, Π_f^h, Π_f^l, Π_m^h, Π_m^l. As will be clear in the data section, the fathers in my estimation sample are all working and the amount of working time is fairly homogeneous over the fathers. This, however is not the case for the mothers and therefore I will let the probability of being a high type mother vary with the mother's socioeconomic status. With this step I attempt to capture that the preference for working of the mother is possibly not random with respect to observable characteristics.[23] The probability functions are of logistic form. For the father: $\frac{\exp(\Psi_{f,0})}{1+\exp(\Psi_{f,0})}$. For the mother: $\frac{\exp(\Psi_{m,0}+\Psi_{m,1}soc1+\Psi_{m,2}soc2)}{1+\exp(\Psi_{m,0}+\Psi_{m,1}soc1+\Psi_{m,2}soc2)}$

The outcome variables are the parental labor supplies and the non-parental childcare choice. Regarding the labor supply, the number of weeks of work is an important outcome variable for the mother in order to measure the total parental leave payment that she gets. For the mother I therefore construct grids of weeks as well as of total hours. For the father it is not important how he distributes the hours over the weeks and thus he has only one grid of total hours. The parents can work for 41 weeks in total with the maximum of 80 hours per week which implies that maximum hours one can work, T, is 3280 hours. Regarding the childcare choice it is modeled as a 0-1 vector and the choice is then multiplied with the chosen mother hours of work of the mother to get total childcare time and the associated cost.

[23] If one was to follow the mothers from when they started accumulating human capital and chose a career path, the preferences that the mothers may have had from birth would most likely have been correlated with occupational choice.

For each combination of the characteristics of the parents in the household, Ω_h and the values of the preference and production function parameters, θ, the model is solved by first finding the non-cooperative Nash equilibrium to obtain the utility of each parents in a potential disagreement situation. Then these indirect utility functions is used as exogenous parameters in the solution of the cooperative Nash problem. The solution to this problem gives the policy functions for parental labor supply, the weeks the mother labor supply is distributed over, and the choice of non-parental childcare or not for each of the combination of characteristics, $m_f(\Omega_h)$, $m_m(\Omega_h)$, $week_m(\Omega_h)$, and $co(\Omega_h)$. These in turn determines the values of child development and consumption.

1.2.4 Model discussion

A couple of points should be noted when it comes to the model. As described in this section the model used for this study is a one period model of parental decision-making in the first period after the birth of the child. While a static model significantly increases the tractability of the model, this approach of course puts a constraint on what can be investigated. First of all, the model cannot be used to investigate how a policy that increases the duration of paid leave affects future wages, career options and future bargaining power. All these potential concerns of the individuals are captured by the preference parameters, as for example the disutility of work. Secondly, unpaid leave may be an important determinant of whether the mother returns to work quickly after the birth of the child or not. If she can take leave and still keep her job, she may be willing to do so even if the leave is unpaid. In a one-period model where there is no continuation value of employment, only the paid leave will be important. Thirdly, to the extent that there are commitment problems in a couple, this may also affect the willingness of the mother to take longer leave. If leave affects the future wage, through for example human capital depreciation, and the consumption of the individual is lower in case of divorce, then the mother would work more when facing a positive probability of divorce in order to insure for such an event.[24] If the probability of divorce is correlated with socioeconomic status or wage, the preference parameters will capture this interesting effect as well. However, the point of this paper is foremost to investigate whether an increase in paid leave directed to the mother may have unintended consequences through changes in the parents' bargaining power. To this aim a one-period model is sufficient.

[24]See Johnson and Skinner (1986) for early empirical evidence of females increasing their labor supply when facing a divorce.

1.3 Institutional background and data

The data I will use in the estimation is obtained from the Millennium Cohort Study, U.K. As I will use the first wave of the study which was carried out in the years 2000-2002 I will first describe the maternity leave payment in place at this point in time. After this I will go through the sample selection and the variable construction.

1.3.1 Maternity leave payments in the UK 2000-2002

In the following section I will briefly describe the leave system in place in the UK at the time of the survey with particular focus on the maternity leave payments.[25].

With the Employment Relations Act 1999 the period of ordinary maternity leave (OML) was extended from 14 to 18 weeks. All women who became pregnant while in employment were statutorily entitled to OML.[26] If the woman had one year of continuous service before the 11th week of the expected due date, she was entitled to an additional period of maternity leave. This period stretched from the end of OML to 29 weeks after child birth.

The Employment Rights Act 1999 states that all parents (including fathers) in continuous service of one year before birth are eligible for 13 weeks of leave. This is normally referred to as statutory parental leave.

Regarding the leave payments, a mother had the right to receive up to 18 weeks of Statutory Maternity Pay (SMP). To qualify for SMP a woman must have been continuously employed by the same employer for at least 26 weeks up to the qualifying week which was the 15th week before the expected birth of the child. Moreover, her average earnings must have equaled the lower earnings limit for National Insurance Contributions. The SMP was separated into two parts: 6 weeks in which the mother received 90 percent of her average weekly earnings from the year before without upper limit and 12 weeks in which she received a flat rate of 62 pounds (Smeaton and Marsh, 2005, p.38). Mothers that did not qualify for the SMP could get maternity allowance (MA). To qualify for MA the mother must have been employed 26 of the 66 weeks leading up to the expected week of birth and had paid 26 weeks NIC during the same period.

When it comes to the fathers, they had the right to take time off from work with the statutory parental leave, but as this leave was unpaid they did not have any legal right to

[25]References from Hudson, Lissenburgh and Dikmen (2002), Maternity and Paternity Rights in Britain: Survey of Parents 2002; and legislation.gov.uk (The Maternity and Parental Leave) unless otherwise stated.

[26]With the condition that the woman notified her employer about the leave-taking a certain number of weeks before the expected birth.

paid leave at this time. However different kinds of employment contracts could offer leave payments also for fathers.

1.3.2 Data

For the purpose of this paper I use data from the UK Millennium Cohort Study (MSC). This is a longitudinal survey of around 19,000 children that were born in the UK at the beginning of the millennium and it currently exists in four waves. The first survey took place in 2001 and 2002 when the children were around nine months. The MCS oversamples families living in areas of child poverty in Northern Ireland, Scotland, and Wales as well as families in areas with high ethnic minority populations in England.

As I focus only on the first period after the birth of the child I will only use information from the first wave. The first survey records information related to the pregnancy, the birth, the health of the mother and the child, the socioeconomic status of the family, living conditions, etc. The main advantages of this data set for the current study is that it contains fairly detailed information on the labor supply of the parents, and especially the mother, in this period and on the parental leave taken by the mother. The main respondent is most often the woman in the household and she answers the more detailed questions about the child, the pregnancy, and the time after birth, but there is also a shorter interview conducted with the partner, if there is one.

As was mentioned in the model section, given their preferences the parents make decisions based on the hourly wage of each of the parents, the education of the mother, the leave entitlements of the mother, and the birth weight of the child. The outcome variables are the total labor supply in hours of each of the parents as well as the number of weeks at home of the mother in this period, and whether the parents used childcare or not. In turn these variables determine income, child development, childcare costs and consumption. In the following I will explain how the characteristics and the outcome variables are constructed from the data at hand.

Sample and variable construction

In this study I only use couples where the main respondent of the survey is the birth mother of the child. Moreover I exclude same sex couples and couples where the partner respondent started living with the mother after the birth of the child. Both parents should be at least 16 years old and not recorded as students, retired, or in a labor market program. Further I exclude couples with more than one child as the model does not feature any

trade-off between different children when it comes to investments. Couples for which there was relevant information lacking on important characteristics or outcome variables are also excluded. Partner respondents who are not currently registered as working and mothers who stated that they never worked are excluded as these individuals possibly have an unobservable trait that is hard to capture in the estimation. As the children in the survey were between 8 and 12 months there are big development differences due to the differences in age. Although I control for the age of the child in weeks throughout, I am not sure that I would capture all related effects and I therefore include only the children of age 9-10 months. I also drop children with stated birth weight under 1 kilo. The final sample contains 3290 couples.

In the following I will refer to the partner respondent as the father.

The wage. The wage rate is an important variable for this study as it is a measure of the individuals' productivity in the labor market. The survey does not provide information on the exact hourly wage and this has to be computed from information given on income and hours worked.[27] The income is unobserved for individuals who do not work or that for some other reason did not provide any labor income information. A wage equation is therefore estimated and used to predict the wage of the individuals. The information used in the estimation of the wage equation is educational attainment, age, socioeconomic status, and race.[28] I exclude individuals who work less than 5 hours a week and who has an imputed hourly wage of less than 0.5 pounds and more than 100 pounds from the estimation. In order to reduce problems of measurement error in the observed wages I use the predicted wage for all individuals as measure of potential labor market productivity.

Eligibility for and computation of leave payments. The mother's labor supply in the previous period determines how much income she can receive per week in maternity leave as explained above. Unfortunately it is not possible to determine the labor market experience of the parents as there is very little information on labor supply from periods before the survey. In order to get a grasp of the distribution of payment entitlement in the sample it is possible to use information from the survey on whether the mothers worked during pregnancy as no work during pregnancy implies that the mother is not entitled to any maternity leave. Further there is information on whether she went on maternity leave after the birth of the child, if she received payments during this period, and if she did, what type of payments.

[27] Importantly I assume here that the income has not changed much for mothers before and after the birth as the estimated wage could be based on income after birth (this is not clear from the survey).

[28] Unfortunately it is not possible to determine the labor market experience of the parents as there is very little information on labor supply from periods before the survey.

74 percent of the mothers stated that they took some kind of leave and almost 90 percent of these mothers received SMP or SMP plus additional pay. The additional pay is most often payments that the mothers receive from their employer. As there is no information on how much the mothers received in additional pay it will be disregarded and the assumption made is that all mothers who received SMP and money in addition to SMP in fact had SMP.[29] Mothers who did not work during pregnancy, who did not take maternity leave and mothers who took maternity leave, but did not receive any payments are assumed not to be eligible for any leave payments. Thus, the mothers in the sample are either assumed not to be entitled to leave payments or they are assumed to be entitled and receive SMP benefits. As described in Section 1.3.1, when receiving the SMP the mother gets a fraction of 0.9 of her average income per week in the previous period for the first 6 weeks and then there are the additional 12 weeks in which the mother receives 62 pounds per week. The survey does not provide any information neither on the income nor on the labor supply from the year before. Therefore I try to approximate the previous labor supply and use this to calculate an approximate (weekly) income from the previous year using the wage of the mother. In order to do this I use information on the average hours worked in the period after birth as well as information on whether the mother changed work hours after birth. For the mothers who stated that they did not change work hours I assume that the current hours are also the hours from the year before. For the mothers who did change hours I check whether they increased or decreased the hours. If they increased them and they are now working full-time (30 hours and above) I assume they used to work part-time before birth. If they stated that they decreased the hours and they work part-time now, I assume that they used to work full-time.

To calculate the leave payment in the model I follow the payment scheme in place in the UK described in the former section. If the mother is home for 6 weeks or less she gets the payment for these 6 weeks, which is $0.9 \cdot w_m \cdot M_m \cdot week_m$, where $week_m$ is weeks at home. If she is home more than 6 weeks, but less than 18 she gets $0.9 \cdot w_m \cdot M_m \cdot 6 + 62 \cdot week_m$. If she is at home more than 18 weeks she get the full amount $0.9 \cdot w_m \cdot M_m \cdot 6 + 62 \cdot 12$.

Mother educational attainment. The mother is assumed to have a high education if she has at least obtained a diploma in higher education.

[29] It should be possible to obtain information on average additional pay for different types of mothers from other sources, but for the moment this is not included.

Socioeconomic status. The MCS provides derived variables with the National Statistics Socio-Economic Classification (NS-SEC) for the respondent grouped into different numbers of categories. I have chosen the measure in seven levels: High managerial, administrative, and professional occupations; Low managerial, administrative and professional occupations; Intermediate occupations; small employers and own account workers; lower supervisory and technical occupations; semi-routine occupations; routine occupations. Individuals who have never worked or are long-term unemployed can be said to constitute an eight category. This variable is further divided into 3 categories with the two highest socioeconomic classes in one, the two second highest in one, and the three lowest in one.

Birth weight. The birth weight in the estimation is constructed as a category variable. Either the child has a low, a normal, or a high birth weight. The baby is considered of low birth weight if its birth weight was lower than 2.7 kilograms and of high birth weight if the weight was higher than 4 kilogram.

Child development The child development at around nine months is constructed by development variables provided in the survey. More specifically, the mother answers questions about how well the child does in terms of gross motor development, fine motor development, and communicative development. For each of the activities under the different labels the mother states whether the child does it often, once or twice, or have not done this yet. The answers are coded with numbers from 1 to 3 where 1 is not yet and 3 is often. A summary of these development variables is provided in Table 1.1. In order to construct a development index I use principal component analysis. The idea is that if there is correlation among several variables there may be an underlying variable that are common to all the variables and principal component analysis is one way of getting to this variable. In this case, the development indicators have the unobserved variable 'child development' in common. As the indicator variables are discrete I use a version of principal component analysis based on the polychoric correlation, which is better suited to the discrete variable case (see Angeles and Kolenikov, 2009). It is estimated using maximum likelihood. There are two variables that the maximum likelihood estimation cannot handle together in this case: sit and walk. This is potentially because the ML estimation estimates them to have a correlation of one. The variable 'sit' is therefore dropped from the estimation as I consider walk to be more informative of the development level. In order to have a non-negative index as this would complicate the utility calculations I have added the minimum value of the index first constructed to the measure of child quality for all couples. In the appendix I show that the

development measure at nine months is strongly linked to cognitive outcomes at age 7 which is the latest age at which the study is currently available.

Table 1.1: Development indicators

Variable	Mean	Std
Gross motor development		
Sits up	2.98	0.18
Stands up holding on	2.58	0.74
Walks a few steps	1.15	0.46
Fine motor development		
Hands together	2.77	0.56
Grabs objects	2.99	0.09
Holds small objects	2.88	0.40
Passes a toy	2.96	0.24
Communicative development		
Gives a toy	2.49	0.69
Waves bye-bye	2.09	0.81
Extends arms	2.82	0.45
Nods for yes	1.18	0.50

Labor supply. The period of interest for this study is set to be 41 weeks after birth for all couples. Further I assume that both parents can work a maximum of 80 hours per week on average which implies a value of T of 3280 hours for each parent. Regarding labor supply in this period it is possible to obtain information on what week the mothers returned to the labor market and how many hours they work on average each week. From this information I construct measures of weeks at home in this period, of hours worked per week if working, and a measure of total hours worked. Regarding the fathers there is a question on whether they took any leave in order to stay at home with the child, which around 80 percent of the fathers in the sample responded positively to. However, there is no information on the actual amount of time spent at home. As other studies made around this time report that most of the fathers who took leave took about one week I will assume that the leave taken by the fathers is so small that is can be disregarded.[30] A measure of the hours worked of the fathers is obtained by assuming that fathers who are recorded as employed worked all 41 weeks at the average number of hours per week that is recorded in the survey.

[30]Maternity and Paternity Rights in Britain: Survey of Parents 2002.

The actual time spent with the child is rarely observed in the datasets used in similar studies and most studies use parental (maternal) employment and/or the use of childcare as a proxy for the time spent with the child.[31] The biggest issue with this measure is that one hour less at work while not using childcare does not necessarily imply that the parent spends one hour more with the child. Actually, findings from time diary studies indicate that mothers who work more do not seem to care less for their child, instead they have less leisure (e.g. Blau and Currie, 2006). In the MCS there is no information on parental childcare and leisure and this may be a problem for the identification of the production function parameters. However, as the period considered covers the first nine months of the child's life, the child is likely to need constant attention and there may not be that much possibility of leisure time when not working, at least not for the mothers. When it comes to the fathers, the estimate of the productivity in child quality production is more likely to be biased. Unless father time with the child is directly harmful for the child's development, the productivity parameter of the father would be biased downards.

Income. The income used for this study is the labor income of the parents in this period and is calculated as the product of the wage and the labor supply plus the income from the maternity leave payments. Unfortunately there is no information on the exact income from other sources in the MCS.

Non-parental childcare. In the survey there is a question on what type of childcare is used when the mother is at work. If the person who takes care of the child while the mother is at work is not one of the parents the parents are assumed to use this childcare for the total working time of the mother during the period. One problem with this approach is that different types of non-parental care may be of different quality.[32] If low income households, for example, cannot afford care of high quality, the parents in this family may be more inclined to stay at home with the child for this reason and not because they have lower wages, lower preferences for work, or because they are less productive in child quality production.[33] Thus, the parameter estimates of the model could be biased for this reason. The MCS does include information that would allow for an estimation of productivity parameters for different types of care. The problem is that it would make the model very large given all

[31] An exception is Del Boca, Flinn, and Wiswall (2011) who are able to distinguish between what they call "active" and "passive" parental time in the child supplement to the PSID.

[32] See for example the literature survey by Brilli, Del Boca and Monfardini, 2013.

[33] As mentioned in the model section, the productivity in child quality production is only allowed to differ between mothers.

other variables already included. An addition of this information is left for future reasearch.

Other information. Apart from the variables the parents decisions are based on I also collect other variables used in, for example the child development regressions. These are the age of the parents, father educational information, father socioeconomic status, age of the child in weeks, the sex of the child, the race of the child, and regional dummies.

1.3.3 Descriptive statistics

Table 1.2 shows descriptive statistics for the characteristics, for the choice variables as well as for other variables that may be of interest. The left panel presents the descriptive statistics of the sample and the right panel shows the same statistic for all the households in the MCS.

In the sample the mothers are slightly younger than the fathers with an average age of 29 years as compared to around 32. More mothers than fathers have a higher education. The wage the father could earn in the labor market is about half a pound higher than the wage of the mother on average. As discussed above, almost 80 percent of the sample mothers have had paid maternity leave in this period and around half the sample has used non-parental childcare. 66 percent of the mothers are working as well as all the fathers as the latter was a sample inclusion criteria. Not surprisingly the mothers work on average less than the fathers in this period, conditional on working.

It is of interest to see how the descriptive statistics of the sample used in the estimations compares to those of all the observations in the MCS (including single households). With appropriate weighting these could be representative of the population and these statistics are shown in the right panel of Table 1.2. We can see that the parents are younger on average in the sample than among all the observations which is not strange since I include only parents with one child. Further, in the sample there is a higher share of married couples and the share of mothers and fathers with at least a diploma in higher education and with a higher socioeconomic status is greater in the sample. Further, in the sample a higher share of mothers got maternity leave payments, more families choose to use non-parental childcare, and a higher share of the parents are working. Finally the income in the sample is higher than in the whole population. The result of the comparison is not surprising considering the sample inclusion criteria of only couples and that only couples with the father working in this period are included.

Table 1.2: Parental characteristics and outcomes sample and all households

	Sample				All households			
Variable	**Mean**	**Std**	**Min**	**Max**	**Mean**	**Std**	**Min**	**Max**
Married	0.69	0.46	0	1	0.59	0.49	0	1
Age mother	29.30	5.11	16	46	29.47	5.84	14	64
Age father	31.77	5.66	17	58	32.82	5.94	16	69
Higher education mother	0.41	0.49	0	1	0.28	0.45	0	1
Higher education father	0.38	0.49	0	1	0.27	0.44	0	1
High socio-econ mother	0.47	0.50	0	1	0.32	0.46	0	1
High socio-econ father	0.54	0.50	0	1	0.39	0.49	0	1
Wage mother	7.05	1.91	3.59	14.78	6.50	1.87	2.71	14.78
Wage father	7.44	2.02	3.20	12.27	7.20	2.10	0.66	12.29
Paid maternity leave	0.79	0.41	0	1	0.52	0.50	0	1
Non-parental childcare	0.52	0.50	0	1	0.33	0.47	0	1
Working mother	0.66	0.47	0	1	0.46	0.50	0	1
Working father	1.00	0.00	1	1	0.90	0.30	0	1
Hours of work mother	537.30	338.16	10	1925.00	506.22	337.56	2	2337.00
Hours of work father	1934.87	480.65	400	3280.00	1886.52	440.68	41	5002.00
Income	18488.36	6823.22	1491	51690.95	12805.15	8933.27	0	67607.82
Low birth weight	0.09	0.29	0	1	0.10	0.31	0	1
High birth weight	0.10	0.30	0	1	0.11	0.32	0	1
Child development	8.27	0.83	2	10	8.19	0.98	0	10

1.4 Estimation

The parameters of the model are estimated using a method of moments approach. In this approach the structural parameters of the model are chosen so as to make the moments generated by the model as close as possible to their empirical counterparts. In the following, θ denotes the vector of model parameters and $\mu(\theta)$ denotes the vector of theoretical moments generated from the model with the structural parameters θ. μ_d are the same moments obtained from the data. At the true value of θ $E(\mu_d - \mu(\theta)) = 0$.

$$\theta^* = \underbrace{\arg\min}_{\theta}(\mu_d - \mu(\theta))' W^{-1} (\mu_d - \mu(\theta)) \tag{1.16}$$

where W is a weighting matrix, here set to be the variance covariance matrix of the data moments.

The parameters to be estimated are the preference parameters, the production function parameters, the unobserved heterogeneity and the probabilities of being of a high or a low type for each parent. In total there are 26 parameters to be estimated.

1.4.1 Moments

The moments used to identify the model are all listed in Table 1.3. There are first and second order moments as the mean and the standard deviation of the outcome variables: labor supply of mother and father respectively, consumption, and child development. To be able to say something more about the differences in preferences I use the fact that with the current model, wage should be linked to bargaining power and attempt to fit the wage coefficient in OLS regressions of hours of work of respective parent on the wages of both parents. In addition I use average hours worked over wage bins (in total 12 bin-moments). Further, in order to identify the child development production function I try to fit the coefficients from an OLS regression of child development on the following variables: Dummies for high and low birth weight; a dummy for whether the mother has high education; hours of work of the mother and the father; an interaction term of their hours of work; an interaction term of high education and hours of work of the mother; as well as a dummy for whether the child is in non-parental care or not. All regressions are purged of potential regional effects and in the child development regression I control for other variables usually thought to be linked to child development. To identify the parameters of the unobserved heterogeneity in preferences I use the standard deviation of the residuals from regressions of the hours of work of the parents on the wages and the education level of the mother. To identify how the unobserved preferences of the parents are linked to child development I first regress child development on wages and the education dummy for the mother. Then I use the standard deviation of the residual and regress that on the residuals from the two former regressions and try to fit the coefficients. Importantly, the unobserved heterogeneity in preferences for work affects both the time spent at home of the parents as well as the initial ability of the child. The identification of their contribution into the initial ability, β_m and β_f as well as of the substitution parameter, ϕ is possible because I assume that k_m, k_f, β_m and β_f are all positive. Therefore, a_p contributes positively to the initial ability, but negatively to the production of child quality through the time at home of the parents.

As mentioned, the probability that the mother has high preferences for working is allowed to be associated with her socioeconomic status. To identify to what extent I run a regression of hours of work of the mother on parental wages, the dummy for high education, and dummies for the socioeconomic status. The moments are the coefficients of the socioeconomic status. I also try to fit a measure of correlation in the hours of the parents which is constructed by regressing the de-meaned work hours of the mother on the de-meaned work hours of the father.[34]

[34]I use the de-meaned hours and not the correlation because with Stata it is easier to obtain the standard

Table 1.3: Moments for estimation

Moments
Preference parameters
Average consumption
St dev consumption
Average child development
St dev child development
Average hours of work father
St dev hours of work father
Average hours of work mother
St dev hours of work mother
OLS regression of hours of work father on wages of mother and father
OLS regression of hours of work mother on wages of mother and father
Average hours of mother and father over father wage bins (6 in total)
Average hours of mother and father over mother wage bins (6 in total)
Child development
OLS regression of child development on birth weight dummies, mother high education, hours of work of the parents, interaction hours of work, interaction terms of mother high education and hours of work mother, dummy variable for non-parental care
Average child development over income bins (4 in total)
Unobserved heterogeneity and type distribution
Standard deviation of residual of OLS regression of father hours on wages of parents and qualification level of mother
Standard deviation of residual of OLS regression of mother hours on wages of parents and qualification level of mother
Standard deviation of residual of OLS regression of child development on wages of parents and qualification level of mother
OLS regression of mother hours on wages and the measure of maternity leave entitlement
Other
OLS regression of demeaned mother hours on demeaned father hours

The regressions of hours on wages include regional dummies.

The regression of child development also includes controls for birth weight, sex, age in weeks, mother and father age, mother and father qualification levels, race, and region.

Finally, in order to say something about the effect and curvature of income in child development I try to fit the model to average child development over income bins.[35]

error of the estimate in the former case.

[35]Since the MCS oversamples families living in areas of child poverty in Northern Ireland, Scotland, and Wales as well as families in areas with high ethnic minority populations in England I calculate all the

1.4.2 The simulated moments

The policy functions given by the solution to the model $m_f^C(\Omega_h;\theta)$, $m_f^C(\Omega_h;\theta)$, $week_m^C(\Omega_h;\theta)$, and $d_{co}^C(\Omega_h;\theta)$ describe the optimal behavior of the parents for each combination of characteristics with the current parameter values. The policy functions are used to simulate the outcomes for each couple by linear interpolation based on the characteristics of each couple in the sample. For each couple of parents I simulate four replicas, one for each parental combination of preference heterogeneity with the corresponding probability of observing this couple combination. This provides me with a cross sectional data set from which I construct the simulated moments. In the moment calculations the simulated data is weighted using the sample weights provided in the MCS as well as the type distribution in the simulated sample. The simulated moments are then matched with the moments from the observed data.

1.4.3 Results

Model fit

When it comes to the model fit, the model does very well. In Table 1.4 the model moments are compared to the data moments for all moments except for hours of mother and father over wage bins and average child quality over income bins. This latter collection of moments is displayed in Figure 1.4 and Figure 1.5. Of the moments displayed in the table the model manages to fit around 82 percent in the sense of being within the 95 percent confidence interval and it is close to fit many of the other moments. The only moment that is not that close is the standard deviation of child development.

Turning to Figure 1.4 we can see that the model fits the parental hours of work over wage bins fairly well and from Figure 1.5 we can see that the fit of child development over income bins is also good.

Parameter estimates

The final parameter estimates are shown in Table 1.5.

Preferences From Table 1.5 we can see that the curvature parameters of consumption are estimated to be quite low, and the curvature of child development is slightly higher. This means that over the sample, the parents would like to keep a fairly equal level of consumption

moments using the survey weights for UK analyses.

Table 1.4: Moment fit

Moment	Model	Data	St errror	CI low	CI up
Average income	18715.2189	18487.9130	251.0223	17994.3504	18981.4757
Std income	6710.1289	6822.4390	141.3779	6545.2080	7099.3989
Average child development	6.2137	6.2276	0.0181	6.1920	6.2633
Std Child development	0.2600	0.8457	0.0159	0.8145	0.8770
Average hours of work father	1936.6813	1934.7975	13.1208	1908.9993	1960.5957
Std hours of work father	480.2038	480.6180	9.7329	461.5323	499.6844
Average hours of work mother	354.4075	357.2191	8.7126	340.0883	374.3499
Std hours of work mother	377.0307	374.6356	5.7901	363.2798	385.9765
Average hours in non-parental care	286.1689	290.2623	8.7809	272.9972	307.5275
Demeaned hours regression	-0.0211	-0.0248	0.0164	-0.0571	0.0075
Coefficients from OLS regression of work hours on wages					
Father hours father wage	51.7829	40.5396	6.4403	27.8766	53.2026
Father hours mother wage	-40.6260	-16.6165	10.7384	-37.7304	4.4974
Mother hours father wage	-50.7680	-50.8457	4.0163	-58.7425	-42.9489
Mother hours mother wage	61.7901	33.3025	9.9432	13.7520	52.8530
Components of child development: Coefficients from OLS regression of child development					
Low birth weight	-0.6811	-0.6177	0.0956	-0.8057	-0.4298
High birth weight	0.1383	0.1660	0.0465	0.0745	0.2574
High education mother	0.0449	0.1151	0.0501	0.0166	0.2137
Interaction education hours mother	-0.0046	-0.0429	0.0154	-0.0731	-0.0127
Hours of work father	0.0214	0.0286	0.0134	0.0023	0.0549
Hours of work mother	0.0241	0.0925	0.0398	0.0142	0.1708
Interaction hours father mother	-0.0081	-0.0055	0.0036	-0.0127	0.0016
Dummy non-parental care	0.2545	0.0641	0.0403	-0.0152	0.1433
Unobserved heterogeneity					
Std residual father	472.5232	474.8834	9.4068	456.4369	493.3109
Std residual mother	362.2059	358.2566	5.7089	347.0603	369.4386
Reg child residual father residual	0.0209	0.0466	0.032585	-0.0175	0.1107
Reg child residual mother residual	0.0450	0.1643	0.045165	0.0755	0.2531
Socioecon status middle	54.8789	46.7047	17.742700	11.8188	81.5906
Socioecon status high	56.5254	130.7771	31.891877	68.0710	193.4833

Note1: To better see the regression coefficients of regression child residual on parental residual, the coefficients from the regressions of a child residual on parent residuals are multiplied by 1000.

and child development and those who start out with a high initial child development (q_0) may spend a little less resources on child development than on consumption, all else equal. Further, the weight on child development is estimated to be higher for the mother than for the father (α_i).

Turning to the parents with high and low preferences from working we can see that around 52 percent of the fathers are of the high type and 63 percent of the mothers. The low

Figure 1.4: Model fit: Average hours worked over wage categories

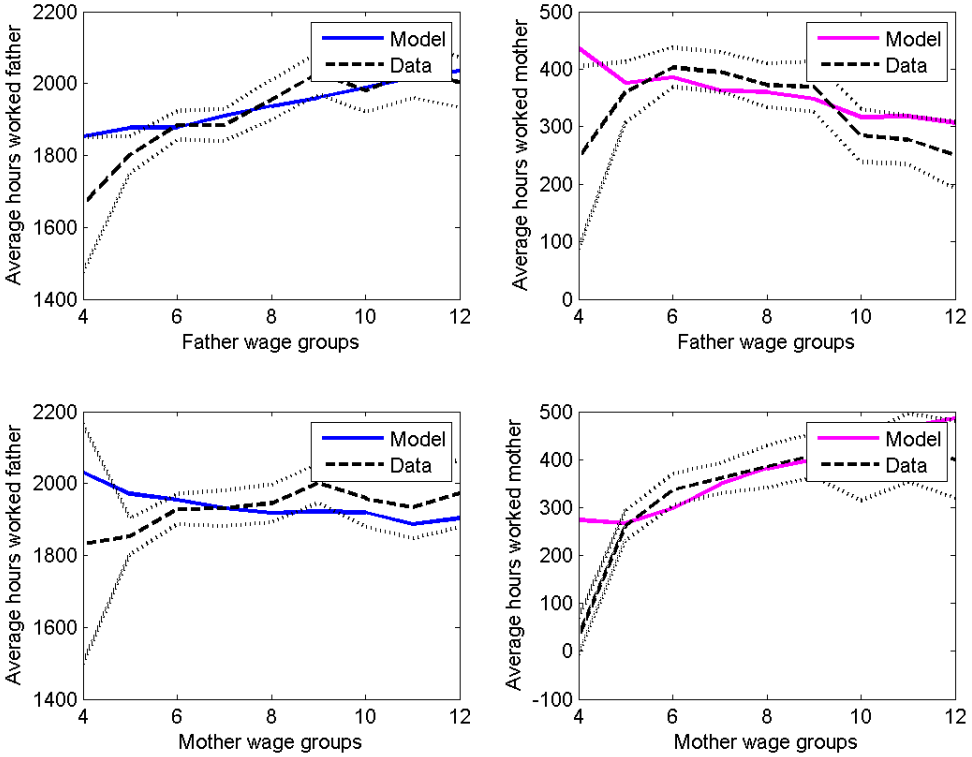


Figure 1.5: Model fit: Average child quality over income bins

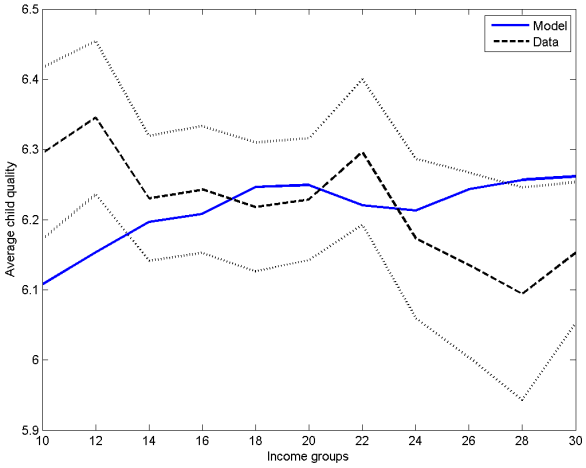


Table 1.5: Parameter estimates

Preferences		*Initial ability*	
δ_c	0.0522	$q_0(\beta_0)$	1.3235
δ_q	0.2807	$q_0(\beta_m)$	0.4504
α_f	4.5313	$q_0(\beta_f)$	0.9022
α_m	13.1843	$q_0(\beta_m, \beta_f)$	1.3526
$\nu_{f,l}$	-0.0000	$q_0(\beta^l_{bw})$	-1.0431
$\nu_{f,h}$	0.0004	$q_0(\beta^h_{bw})$	0.2705
$\nu_{m,l}$	-0.0061	$q_0(\beta_e)$	0.0000
$\nu_{m,h}$	0.0033	*Production function*	
ν_{mf}	0.0000	κ_f	0.0079
p_{cc}	1.0517	$\kappa_m(0)$	0.0227
		$\kappa_m(1)$	0.0230
		ϕ	0.5454
Type proportions		κ_i	0.0000
Π_f	0.5201	ϕ_i	0.6574
Π_m	0.6307	κ_o	0.0001

type parents get disutility from work (ν_i), whereas the high type parents enjoy working. This result may seem strange. In theoretical work it is most often assumed that individuals dislike work and this is also the findings in estimations (e.g. Eckstein and Wolpin, 1989; Bernal, 2008). However, it is important to have in mind that the current model is a one-period model, which means that the preference parameters will capture also future considerations. If absence from the labor market will have a negative impact on wages and the career in general and this is a concern of the individuals the preference parameters will pick this up.[36] Moreover, the model is estimated in a very specific context; the maternity leave period. As the Survey of Parents 2005 (p. 34) reports, the most common reason for the mothers' return to work after the birth of the child was actually a desire to get back to work. It may also be noted that the preference difference between high and low type fathers is smaller than that between high and low type mothers, which could be expected given the sample restriction on working fathers.

The average cost of non-parental childcare is estimated to be around 1 pound an hour. That the price is estimated to be this low is not surprising as non-parental childcare is

[36] Adda, Dustmann, and Stevens (2011) estimates a model featuring both fertility and career choices to investigate the cost of fertility and its sources. They find that atrophy and lower accumulation of human capital due to absence from the labor market constitute an important part of the cost, and especially for women in more abstract occupations who tend to have a steeper wage profile.

assumed to be all childcare not carried out by the parents. As such the definition also includes care by grandparents or other relatives which may be free, at least in monetary terms.

Production function The estimates of the production function show that low birth weight decreases the initial ability of around 1 unit and high birth weight increases the initial ability of around 0.27 units. Having a mother with higher education increases the initial ability slightly. β_m and β_f describe the contribution of the unobserved part of the mother and father preferences on the initial ability respectively. Assuming that the education is low and the birth weight normal, the initial ability from having a high type mother is 1.77 and from having a high type father it is 2.23. If both parents are of a high type the initial child ability is 2.68. In sum, it seems like the father type is somewhat more important for the initial development of the child.

Table 1.6: Initial child development

Normal birth weight and low mother education	
Combinations	Initial child development
Father low, mother low	1.32
Father low, mother high	1.77
Father high, mother low	2.23
Father high, mother high	2.68

When it comes to the other factors determining the child development in this period the parameter values are very low. This is the result of the low average value of child development and the input variables being defined in hours and pounds. We can see that one unit of mother time spent not working (κ_m) contributes more than twice than that of the father (κ_f).[37] The reason could be for example breastfeeding. Further, it also seems like the quality of time not spent working is slightly higher for high qualified mothers ($\kappa_m(1)$) than for low qualified ($\kappa_m(0)$), about 1 percent. The parents are estimated to be somewhat complementary in their time spent with the child. Income and non-parental care have a very small unit effect on child development and the effect of income is a marginally decreasing.

[37]The findings of Del Bocca, Flinn and Wiswall (2010) are in line with the current results as they indicate that both the "active" and the "passive" time that the mother spends with the child when the child is about 2 years old are more productive than the time spent of the father. Also the results of Cherchye, De Rock and Vermeulen (2012) point in the direction of mothers' time with children being more productive than fathers' time.

That current income (similar to the consumption here) does not seem to matter for child outcomes is not an uncommon result. For example, Heckman and Carneiro (2003) argue that it is the permanent income that matters rather than the current one. They show that once they control for test scores or permanent income, the association between income and child outcomes disappears. The result is supported by findings of Bernal (2008): income does not matter for child outcomes when she controls for the mother's education and test scores.[38]

1.5 An increase in payment generosity

In this section I try to answer how an increase in paid maternity leave policies affects the parental decision-making, child development, and the distribution of resources within the household. In order to do this I use the estimated model to study how the behavior and the outcomes change with an increase in the generosity of leave payments. An important aspect of the paper is the effect of the policy change on the bargaining positions of the parents and how this affects the distribution of well-being within the household. To investigate this I compare the outcome in a case when the bargaining power changes with the policy change to the outcome in the case when the bargaining positions are held fixed.

The policy change considered is designed to be similar to a policy change that took place in 2003. As mentioned in Section 1.3.2, the relevant payments for this study are the SMP payments. In 2003 they changed in the following way: The duration of the leave payments was extended from 18 weeks to 26 weeks by increasing the number of weeks of flat rate payments from 12 to 20; and the payments in the flat rate period were set to 100 pounds rather than 62 pounds per week.[39]

To get a better picture of the changes as a result of the policy Table 1.7 shows the baseline prediction (i.e. the parental behavior under the original policy), the prediction under the new policy, and the average percentage change between the cases for the main outcome variables: the total hours of work of the mother and the father in the first 41 weeks of a child's life; child development by the end of the period; consumption; and non-parental childcare during

[38]However, two "natural experiment" studies go in a different direction. Dahl and Lochner (2008) exploit variation in the Earned Income Tax Credit as potentially exogenous variation in income and they find that income seems to have a positive effect on a child's academic achievement. Løken (2010) uses the so called oil boom in Norway as an instrument for an increase in permanent income. She finds no effects on educational attainment.

[39]Unfortunately, the MCS offers no way to check the predictive performance of the model. In order to do this one could potentially look at labor supply data of comparable families before and after the policy change.

the period and the maternity leave payments.

Table 1.7: Results from an increase in the payment generosity

Outcome	**Baseline**	**New policy**	**Percentage change**
Total hours mother	566.62	489.92	-19.92
Total hours father	1936.68	1867.61	-4.11
Child development	6.21	6.31	1.51
Consumption	18413.98	18617.26	1.05
Child care	625.02	540.16	-18.62
Maternity leave payment	2018.99	3274.63	67.18

The hours of work of the mother are conditional on mother working under the original policy, the non-parental childcare variables are conditional on using non-parental care under the original policy, and the maternity leave payment is calculated only for mothers who had payments in the first period.

We can see that the mothers who would work under the original policy decrease the labor supply after the policy change by around 20 percent. When studying the results it is important to keep in mind that around 10 percent of the mothers that work under the original policy are not eligible for the leave payment and will therefore not be affected by the change. Additionally, for many mothers who work before the leave there is some monetary loss from taking the leave instead of working. The average sample wage for the mothers who are predicted to work at the baseline is 7.3. If a mother at this wage would work a full-time week of 40 hours she would earn 280 pounds per week. A benefit payment of 100 is 35 percent of the potential average earnings per week.

Another important aspect is that many mothers who are eligible for the payments stayed home a longer time than what was covered by the original policy. Thus, with the policy change these mothers get money for weeks spent at home that would have been unpaid under the original policy. The income effect of the policy change goes in the direction of decreasing the labor supply. With more income the parents can better afford to take more time off from work in this period and enjoy a bit more child development while still having the same amount of consumption. This effect in combination with the slight complementarities of the parents in the production of child development contributes to the father decreasing his hours of work as well. As we can see in Table 1.7 the average consumption has increased by around 1 percent and the leave payments for the mothers who were eligible before the payment increase have increased by almost 67 percent. As a response to the decrease in the hours of work of the mother the parents also use less non-parental childcare.

Hence it seems like a part of the intended result of the increase in the generosity of leave

payments is obtained. With increase in payment and duration of paid leave the mother is able to spend more time at home taking care of the child. The question is then how much that affects the child development. From row three of Table 1.7 we can see that the average child development in the sample has increased by a modest 1.5 percent as a result of the policy change. An important reason behind this is that the effect of parental time spent with the child is decreasing at the marginal. This implies that the child development is affected more when increasing the leave the first weeks rather than after a couple of months.[40]

1.5.1 At the threat points

In this section I examine how the policy change affects the bargaining position of the parents. As described in Section 1.2, the relative bargaining position in the current set-up is determined by the relative well-being of the parents in the case of a disagreement where they are assumed to act non-cooperatively according to a non-cooperative Nash bargaining game. With the estimated parameters I can simulate also the non-cooperative (unobserved) outcomes before and after the policy change. If one parent gains in utility relative to the other parent as a result of the policy, the policy has changed the bargaining positions of the parents. Such an effect may have important consequences for the well-being within the household. As described in Section 1.2 and in the illustration of Figure 1.3, when the bargaining power of one parent increases relative to the other parent the preferences of the first parent will be weighted higher in the cooperative solution. This in turn will change the resource distribution within the household and this parent would be relatively better off.

Table 1.8 shows the outcomes for the non-cooperative equilibrium. In the first column we find the baseline predictions, in the second column the predictions after the change in policy, and the third column shows the percentage change from baseline. A first thing to note is that at the baseline both the mother and the father work on average much more than in the cooperative solution. This is a result of the nature of the consumption good. As this is private and can only be paid for by private income the parent him/herself needs to work to consume. Moreover, in the set-up of the non-cooperative game the parents do not take the well-being of the other parent into account when they make their decisions. As such they will spend more time on the goods from which they derive utility, not taking into account the effects of this decision on the utility of the other parent. As a result of this time

[40] A non-linear effect in time spent with the child may explain why Carneiro et. al. (2009) find a positive effect of maternity leave on child outcomes, while the other natural experiment studies do not (see Baker and Milligan 2008, 2010, 2011; Dustmann and Schnberg, 2008; Liu and Nordstrm Skans, 2010 and Rasmussen, 2010).

allocation, the child development is lower for non-cooperating couples. The consumption of the mother is much lower than the consumption of the father which is mainly due to her working less and paying all the costs for non-parental childcare. As we can see, the childcare usage is fairly high.

Table 1.8: Changes in the case of no cooperation

Outcome	**Baseline**	**New policy**	**Percentage change**
Total hours mother	1077.72	950.27	-10.49
Total hours father	2458.19	2456.55	-0.06
Child development	5.45	5.55	1.91
Consumption father	18324.07	18309.48	-0.06
Consumption mother	6380.83	6721.95	19.73
Child care	1123.07	992.41	31.51
Maternity leave paymentc	2016.66	3162.34	67.18

Note that the hours of work of the mother are conditional on mother working under the original policy, that the non-parental childcare variables are conditional on using non-parental care under the original policy, and that the maternity leave payment change is calculated only for mothers who had payments in the first period.

From the second and third columns we can see that with the current change in policy the mother would, in the case of disagreement, decrease her working time about 10 percent on average. The father would also decrease his working time because of the complementarity in child development production, but the decrease is very small. This move has a negligible negative effect on his consumption. The average maternity leave payments would increase from about 1026 to about 3162 and the mother increases her consumption of about 20 percent. Due to the decrease in working time of the mother the child development increases by around 1.91 percent from the baseline prediction.

If we compare how much better off the parents are after the policy change we first note that both parents are better off. Measured in consumption units the mother is 830 better off and the father is 1350 better off. However, even if it seems like the father may have gained in utility relative to the mother, one cannot draw this conclusion from the consumption values. Due to the decreasing marginal utility of consumption and the fact that the father consumes much more than the mother under the original policy he needs a bigger amount of consumption than the mother to gain one unit of utility. Another way to look at this is to compare the relative utility of the parents before and after the policy change. The average relative utility of the father to the mother at the baseline is 0.5429 and after the policy change it is 0.5425. Thus mother seems to have improved her bargaining position slightly

relative to the father on average. However, it may be interesting to study this result closer over types. As we noted in Table 1.5, some mothers are estimated to have positive utility from work (63 percent) and one may expect that the results might look different because of this. The relative utility is shown in Table 1.9 and we can see that the high type mothers have a lower bargaining position after the policy change, whereas the low type mothers have a higher bargaining position. As hinted at, this is due to the high type mothers decreasing their labor supply while the fathers can enjoy more child development without sacrificing any consumption or labor supply. Therefore, the father increases his utility relative to the mother in a household where the mother is of a high type.

Table 1.9: Relative utility father to mother no cooperation

	Mother type		
	l	h	**Total**
Baseline	0.5603	0.5327	**0.5429**
New policy	0.5572	0.5339	**0.5425**

1.5.2 Policy change with the bargaining position fixed

In order to check whether the changes in bargaining positions have effects on the household decision-making and the distribution of resources and well-being in the household I compare the case when bargaining power changes to a case when the bargaining power is held fixed. The average results are shown in the top of Table 1.10. In the very left column we have the baseline predictions, the next two columns show the predictions for the policy change when the bargaining power is fixed. The two columns to the far right show the results when the bargaining power changes, which is the case studied above.

As we can see there is not much change in the mother work variables between the cases. When the bargaining positions are fixed she decreases the work hours slightly less than when the bargaining power is allowed to change. The decrease in father hours is almost one percent bigger in the case when the bargaining power is fixed. The change in time distribution together with a small decrease in the amount of childcare used implies that child development increases slightly more when the bargaining positions do not change. The change in consumption goes in the opposite direction and is about half a percent higher with a change in bargaining power.

Based on the findings on bargaining power changes we know that the top of Table 1.10 shows the average effect when some mothers gain in bargaining power (a little less than

Table 1.10: Comparing two cases: fixed vs. changing bargaining positions

		Change in bargaining positions		Bargain positions fixed	
Outcome	**Baseline**	**New policy**	**Percentage change**	**New policy**	**Percentage change**
Total					
Total hours mothera	566.62	489.92	-19.92	499.13	-17.89
Total hours father	1936.68	1867.61	-4.11	1850.36	-5.01
Child development	6.21	6.31	1.51	6.31	1.56
Consumption	18413.98	18617.26	1.05	18524.52	0.67
Child careb	625.02	540.16	-18.62	551.48	-15.95
Maternity leave paymentc	2018.99	3274.63	67.18	3274.61	67.18
Mother low type					
Total hours mothera	18.884	15.716	-22.208	15.716	-22.208
Total hours father	2240.060	2123.071	-5.821	2143.617	-4.820
Child development	6.132	6.206	1.213	6.193	1.003
Consumption	18281.866	18448.086	0.746	18590.124	1.617
Mother high type					
Total hours mothera	576.609	498.560	-19.881	507.938	-17.813
Total hours father	1759.020	1718.024	-3.112	1678.636	-5.128
Child development	6.262	6.368	1.691	6.379	1.882
Consumption	18491.347	18716.322	1.225	18486.101	0.118

Note1: The mother hours of work are conditional on mother working under the original policy

Note2: The non-parental childcare variables are conditional on using non-parental care under the original policy

Note3: The change in leavepayment is calculated only for mothers who had payments in the first period

half of the mothers) while the other part loose. In order to study how the intra-household distribution of resources is affected by the mother gaining or loosing in bargaining power, the lower part of the table divides the sample into households with mothers of low preferences for working and mothers of high. Here we can see that most of the action is in the case of high type mothers. First of all, mothers who prefer to be home would not work much under the original policy and they do not change behavior much in hours in absolute terms. The father would work slightly less when the bargaining power changes, the child development would increase a bit more and the consumption a bit less. This is in line with the preferences of the mother as she puts a higher weight on child development relative to consumption.

If we instead study the bottom of Table 1.10 we still see the interaction of bargaining power and preferences. Remember that in this case it is the bargaining power of the father that increases. When the bargaining power changes the mothers would work less and the fathers would work more (as most fathers do not mind working). The child development is lower when the father gains in bargaining power and consumption is higher. This is in line with the father's preferences.

To get a notion of the changes in well-being for the two cases I calculate the consumption

value of the policy change for the parents. Note that the average consumption under the original policy is about 18400. Not surprisingly both parents gain from the increase in the generosity of leave payments, but for whom the gain is bigger differs between the cases. When the bargaining power changes with the policy the mother is 1470 pounds better off and the father is 960 pounds better off. However, when the bargaining positions are held fixed the mother is 1750 pounds better off and the father is 800 pounds better off. Both parents gain from a policy change, but the mother would on average gain more and the father less when the bargaining power does not change.

1.5.3 Heterogeneous effects

It may not only be of interest to look at the population average when analyzing the effects of a maternity leave policy. If more generous leave payments would have different effects for lower income households than for higher income households it may affect the desirability of the particular policy depending on the preferences of the policy maker.

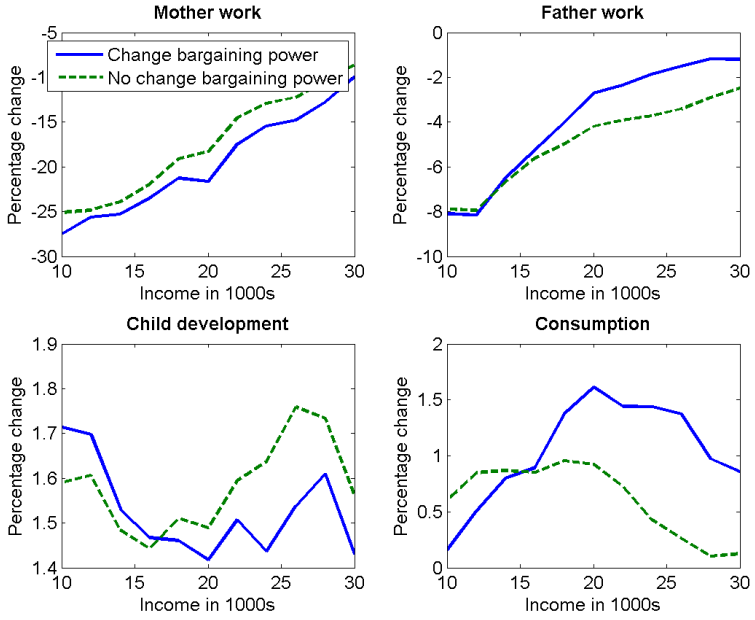
In Figure 1.6 we can study how the behavior of the parents changes with the policy over the labor income. The filled line represents the case when the bargaining power changes and the dashed line represents the case when the bargaining power is held fixed. If we consider the change in bargaining power as the case we would observe in reality with a change in the policy we see that the relative change in mother and father hours is bigger for the low income groups. This is not surprising as the payment is bigger relative to the wage for these households. For the child development the filled line predicts that the relative change in child development is slightly higher for low income groups and that the relative increase in consumption would be higher for high income groups.

If we then consider the situation when the bargaining power is held fixed, we see that the mother would work more and the father less, reflecting the influence of the high-type mothers. The low-type mothers are more prevalent in the lower income groups and the child development actually increases there under a bargaining power change, whereas the opposite is true for the higher income groups. For consumption the pattern is the opposite to that of child development.

Figure 1.7 shows the utility changes of the parents over labor income with and without changes in bargaining power. Essentially, in all the income groups the father would be better off if the bargaining power changes whereas this is only the case for the mother in the lower income groups. This implies that the father has improved his position in the household as a result of the policy in all income groups, except for in the lower income groups.

When studying the changes over income groups it becomes more complex. While the

Figure 1.6: Changes in behavior with and without change in bargaining power

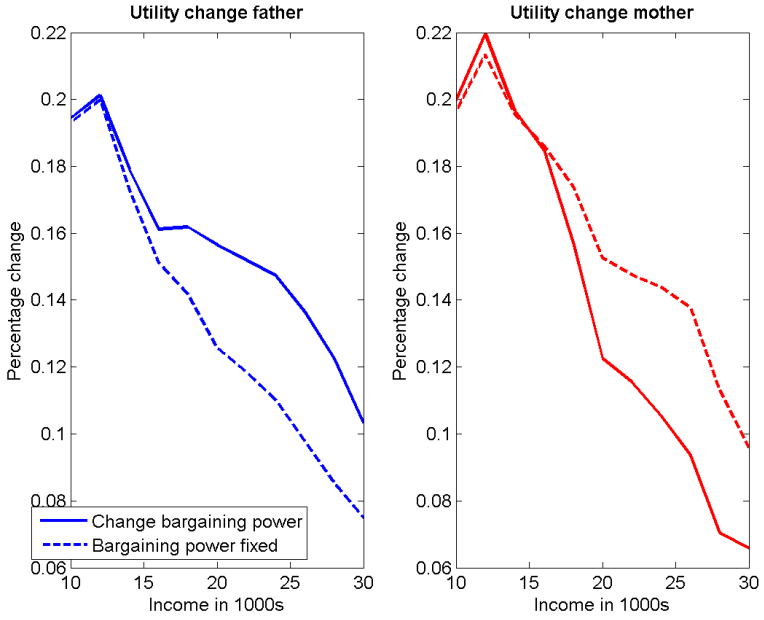


increase in child development is modest, it seems like the children of low income households would gain slightly more relative to the high income households. These households are also the only households where it is not clear that the father has gained in well-being relative to the mother as a result of the policy.

1.6 Conclusion

Generous maternity leave policies are often motivated on the basis of the well-being of the child and of the family, even though it is not clear how these outcomes are actually affected. First of all, existing evidence do not show that neither short nor long run child outcomes are significantly improved by increases in periods of maternity and parental leave. Moreover, while the well-being of the parents on a household level is likely to increase it is not clear how their relative well-being within the household is affected. When investigating effects of policies we might want to consider that a household actually consists of two individuals that may be unequally affected by policy changes. In this paper I look closer at these issues. In particular I estimate a static model of parental decision-making in the first nine months of a child's life where decisions are made through cooperative Nash bargaining and the threat

Figure 1.7: Utility changes with and without change in bargaining power



points, or the bargaining positions, are represented by a non-cooperative Nash equilibrium. The model is used to predict how an increase in maternity leave payments affect parental labor supply, early child development, and the relative well-being of the parents.

The findings suggest that with the increase in payment generosity both mothers and fathers spend less time working in this period, although the decrease in the fathers' labor supply is relatively small. Regarding the effect on child development, despite the decrease in the parental labor supply and an average increase in consumption, the increase in child development is small on average. Interestingly this seems to be due mostly to the decreasing marginal returns to parental time input.

Furthermore, I find that the policy change has an effect on the bargaining positions within the household, but that the effect does not go in the same direction in all households. In households where the mother gets a lot of utility from staying at home, she will gain in bargaining power from the policy change whereas the mothers who would like to work more looses in bargaining power. The main explanation behind the result is a lower labor supply of the mother in the non-cooperative situation which affects the different mothers differently. In the case of the high type mother (who likes work), even though she enjoys an increase in child development and consumption, the father is able to free-ride on her

contribution to child development as he can enjoy it without decreasing his labor supply and consumption. Comparing the parental behavior in the case when the bargaining power is changing with the policy to a case when the bargaining power is held fixed I show that this change in bargaining positions matters and that the results seem to be driven by the mothers with higher preferences for working. When the bargaining power of the father increases, the household shifts resources from time spent with the child to more labor market work of the father, and to consumption. This is in line with the father's preferences which are weighted higher in the household objective function. Consequently, even though both parents are better off from the policy change, on average the father gain in utility relative to the mother.

An analysis over income groups shows that the relative increase in child development is bigger for mothers in the lower income groups, although the differences are small. The reason is mainly a relatively large decrease in the labor supply of the parents. If we look at the utility changes over the income the father gains in utility relative to the mother with the change in bargaining positions, except for in the low income groups.

The conclusion from this paper is that a more generous maternity leave policy provides for better consumption possibilities and facilitates for the parents to stay at home with their children after the birth of the child. However, an increase in the generosity of payments does not seem to have much effect on the child development during the first year. Moreover it affects the bargaining positions of the parents and on average in the sample, the mother would have been better off relative to the father if the policy would not have changed.

1.7 Appendix Chapter 1: The development at 9 months and child cognitive measures at age 7

It is of interest to know whether development measure constructed here is associated with future outcomes. As the Millennium Cohort Study is a longitudinal study with child outcomes measured at each wave it is possible to do this. The survey is currently available in four waves and the fourth wave takes place when the child is around 7 years old. There are three different measures of child cognitive outcomes at age 7 provided by the survey. The first is the BAS word reading score, based on the child reading words out aloud from a card. Here I use the standardized score as outcome variable. The second measure is BAS pattern construction. The performance of the child in this task is based on accuracy and speed when it comes to putting together flat squares or solid cubes with yellow and black patterns on each side. Here I use the age-based t-score provided by the survey. The third measure is a math score based on completing number-based tasks in topics such as numbers, space and measures, and the handling of data. For this measure I use the total score.[41] Table 1.7 shows the results from OLS regressions of these measures on the measure of child development.[42] The results indicate that the development measure at 9 months is a strong predictor of child cognitive development at age 7.

.Child cognitive development at age 7

	(1)	(2)	(3)
	Word reading	Pattern construction	Maths Test Score
Development at 9 months	1.480	1.657	0.323
	(0.691)**	(0.344)***	(0.085)***
Observations	2600	2600	2600

Standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%
Other controls included: Age of baby and parents; birth weight of baby; dummy for female; education of the parents; income and hours of work in the first period; race and region dummies.

[41] For more information regarding the child outcome measures, see Jones and Schoon (2010).

[42] As can be seen there is some attrition in the sample. Importantly, the child development is not correlated with the attrition as the mean in the smaller sample is 8.27, the same mean as for the bigger sample.

Bibliography

[1] Adda, J., Dustmann, C. & Stevens, K. (2011). The career costs of children. CEPR Discussion Paper, No. 8697.

[2] Aizer, A. & Cunha, F. (2012). The production of human capital: endowments, investments, and fertility. NBER Working Paper, No. 18429.

[3] Alessie, R., Crossley, T.F. & Hildebrand, V.A. (2006). Estimating a collective household model with survey data on financial satisfaction. SEDAP Research Paper 161, McMaster University.

[4] Angeles, G. & Kolenikov, S. (2009). Socioeconomic status measurement with discrete proxy variables: is principal component analysis a reliable answer? *Review of Income and Wealth, 55*, 128-165.

[5] Baker, M. & Milligan, K. (2008). Maternal employment, breastfeeding, and health: evidence from maternity leave mandates. *Journal of Health Economics, 27*, 871-887.

[6] Baker, M. & Milligan, K. (2010). Evidence from maternity leave expansions on the impact of maternal care on early child development. *Journal of Human Resources, 45*, 1-32.

[7] Baker, M. & Milligan, K. (2011). Maternity leave and children's cognitive and behavioral development. NBER Working Paper, No. 17105.

[8] Berger, L.M., Hill, J. & Waldfogel, J. (2005). Maternity leave, early maternal employment and child health and development in the U.S. *The Economic Journal, 115*, F29-F47.

[9] Bernal, R. (2008). The effect of maternal employment and child care on children's cognitive development. *International Economic Review, 4*, 1173-1209.

[10] Bernal, R. & Fruttero, A. (2008). Parental Leave Policies, Intra-household Time Allocations and Children's' Human Capital. *Journal of Population Economics, 21*, 779-825.

[11] Bernal, R. & Keane, M.P. (2010). Quasi-structural estimation of a model of childcare choices and child cognitive ability production. *Journal of Econometrics, 156*, 164-189.

[12] Blau, D. & Currie, J. (2006). Pre-school, day care, and after-school care: who's minding the kids? In: E. Hanushek and F. Welch, Editor(s), *Handbook of the Economics of Education* (1163-1278). Volume 2. Elsevier.

[13] Brilli, Y., Del Boca, D. & Monfardini, C. (2013). Child care arrangements: Determinants and consequences (No. 18). Center for Household, Income, Labour and Demographic Economics (CHILD)-CCA.

[14] Brooks-Gunn, J., Han W-J. & Waldfogel, J. (2002). Maternal employment and child cognitive outcomes in the first three years of life: the NICHD study of early child care. *Child Development, 73*, 1052-1072.

[15] Browning, M., Bourguignon F., Chiappori, P.-A., & Lechene, V. (1994). Incomes and outcomes: a structural model of intrahousehold allocation. *Journal of Political Economy, 102*, 1067 - 1096.

[16] Browning, M., Chiappori, P.-A., & Lechene, V. (2006). Collective and unitary models: a clarification. *Review of Economics of the Household, 4*, 5-14.

[17] Carneiro, P., Crawford, C. & Goodman, A. (2007). The impact of early cognitive and non-cognitive skills on later outcomes. Open Access publications from University College London http://eprints.ucl.ac.uk/, University College London.

[18] Carneiro, P., Løken, K. & Salvanes, K. (2009). A flying start? Long Term Consequences of Maternal Time Investments in Children During Their First Year of Life. IZA Discussion Paper, No. 5362.

[19] Chen, Z. & Woolley, F. (2001). A Cournot-Nash model of family decision-making. *Economic Journal, 111*, 722-748.

[20] Cherchye, L., De Rock, B. & Vermeulen, F. (2009). Opening the black box of intrahousehold decision-making: Theory and non-parametric empirical tests of general collective consumption models. *Journal of Political Economy, 117*, 1074-1104.

[21] Cherchye, L., De Rock, B. & Vermeulen, F. (2012). Married with children. A collective labor supply model with detailed time use and intrahousehold expenditure information. *American Economic Review, 102*, 3377-3405.

[22] Currie, J., & Moretti E. (2003). Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *Quarterly Journal of Economics, 118*, 1495-1532.

[23] Dahl, G. & Lochner, L. (2012). The impact of family income on child achievement: evidence from the earned income tax credit. *The American Economic Review, 102*, 1927-1956.

[24] Dahl, G., Løken, K.V., Mogstad, M. & Salvanes, K.V. (2013). What is the case for paid maternity leave? NBER Working Paper, No. 19595.

[25] Del Boca, D. & Flinn, C. (2012a). Endogenous household interaction. *Journal of Econometrics, 166*, 49-65.

[26] Del Boca, D. & Flinn, C. (2012b). Household behavior and the marriage market. Carlo Alberto Notebooks 254, Collegio Carlo Alberto.

[27] Del Boca, D., Flinn, C. & Wiswall, M. (2011). Household choices and child development. Carlo Alberto Notebooks 149, Collegio Carlo Alberto.

[28] Dustmann, C. & Schnberg, U. (2012). Expansions in maternity leave coverage and children's long-term outcomes. *American Economic Journal, 4*, 190-224.

[29] Eckstein, Z. & Wolpin, K. I. (1989). Dynamic labor force participation of married women and endogenous work experience. *The Review of Economic Studies, 56*, 375-390.

[30] Erosa, A., Fuster, L. & Restuccia, D. (2010). A general equilibrium analysis of parental leave policies. *Review of Economic Dynamics, 13*, 742-758.

[31] Han, W-J., Ruhm, C. & Waldfogel, J. (2009). Parental Leave Policies and Parents' Employment and Leave-Taking. *Journal of Policy Analysis and Management, 28*, 29-54.

[32] Gauthier, A.H. (2011). Comparative Family Policy Database, Version 3 [computer file]. Netherlands Interdisciplinary Demographic Institute and Max Planck Institute for Demographic Research (distributors). Retrieved from: www.demogr.mpg.de

[33] Hudson, M., Lissenburgh, S. & Sahin-Dikmen, M. (2002). Maternity and paternity rights in Britain: Survey of parents 2002. Department for Work and Pensions, In-House Report. 131.

[34] Jones, E. M. & Schoon, I. (2010). Cognitive development. In Hansen, K., Jones, E., Joshi, H. & Budge, D. (Eds.), *Millennium Cohort Study Fourth Survey: A user's guide to initial findings* (pp 150-164). 2nd Edition. Center for Longitudinal Studies, Institute of Education, University of London.

[35] Konrad, K. & Lommerud, K. (1995). Family policies with non-cooperative families. *Scandinavian Journal of Economics, 97*, 581-601.

[36] Lechene, V. & Preston, I. (2011). Noncooperative household demand. *Journal of Economic Theory, 146*, 504-527.

[37] Liu, Q. & Nordstrm Skans, O. N. (2010). The duration of paid parental leave and children's scholastic performance. *The B.E. Journal of Economic Analysis & Policy, 10*.

[38] Lundberg, S. & Pollak, R. (1993). Separate spheres bargaining and the marriage market. *Journal of Political Economy, 101*, 988-1010.

[39] Lundberg, S. J., Pollak, R. A. & Wales, T. J. (1997). Do husbands and wives pool their resources? Evidence from the United Kingdom child benefit. *The Journal of Human Resources, 32*, 463-480.

[40] Lundberg, S., Startz, R. & Stillman, S. (2003). The retirement-consumption puzzle: a marital bargaining approach. *Journal of Public Economics, 87*, 1199-1218.

[41] Løken, K. V. (2010). Family income and children's education: Using the Norwegian oil boom as a natural experiment. *Labour Economics, 17*, 118-129.

[42] Manser, M. & Brown, M. (1980). Marriage and household decision-making: A bargaining analysis. *International Economic Review, 21*, 31-44.

[43] McElroy, M. B. & Horney, M. J. (1981). Nash-bargained household decisions: Towards a generalization of the theory of demand. *International Economic Review, 22*, 333-349.

[44] Phipps, S. A. & Burton, P. S (1998). What's mine is yours? The influence on male and female incomes on patterns of household expenditure. *Economica, 65*, 599-613.

[45] Rasmussen, A. W. (2010). Increasing the length of parents' birth related leave: The effect on children's long-term outcomes. *Labour Economics, 17*, 91-100.

[46] Ruhm, C.J. (2000). Parental leave and child health. *Journal of Health Economics, 19*, 931-960.

[47] Ruhm, C.J. (2008). Maternal Employment and Adolescent Development. *Labour Economics, 15*, 958-983.

[48] Ruhm, C.J. & Teague, J.L. (1995). Parental leave policies in Europe and North America. NBER Working Paper, No. 5065.

[49] Smeaton, D. & Marsh, A. (2005). Maternity and paternity rights and benefits: Survey of parents 2005. Department for Work and Pensions, Employment relation research series no. 50.

[50] Vermeulen, F. (2002). Collective household models: principles and main results. *Journal of Economic Surveys, 16*, 533-564.

[51] Vermeulen, F. (2005). And the winner is... An empirical evaluation of unitary and collective labour supply models. *Empirical Economics, 30*, 711-734.

Chapter 2

Co-movements in Household Income Shocks

With Andreas Fagereng and Kjersti Naess Torstensen

2.1 Introduction

The dynamics of earnings and wages at the individual and at the household level are much studied in the literature and more recent papers point to important heterogeneity.[1] One parameter that might be important for the observed patterns, at least in households where there is more than one income earner, is the co-movement of income of the individuals in the household. Heterogeneity in the co-movement of income is important as it could indicate that marriages differ in their risk-sharing benefits. When there are shocks to individual wages spouses could provide insurance through income pooling and through labor supply, but a higher correlation of wage shocks would make both insurance channels less effective. Since labor supply is potentially not perfectly flexible we would expect the income correlation to carry some information also on the second risk-sharing channel.[2] In general, co-movement of income could also indicate to what extent the individuals are substitutes or complements in market and household production as well as tell something about their preferences to spend leisure time together.

[1] See for example the series of papers by Moffitt and Gottschalk (1994; 1995; 2002; and 2006), and Meghir and Pistaferri (2004) as well as Blundell, Graber, and Mogstad (2014) for papers emphasizing heterogeneity.

[2] That spousal income and labor supply are important for spousal insurance is indicated by several studies (e.g. Kotlikoff and Spivak, 1981; Lundberg, 1985; Cullen and Gruber, 2000; Stephens, 2002; Blundell, Pistaferri, and Saporta-Eksten, 2012; and Ortiguera and Siassi, 2013).

With this paper we contribute to the scant literature studying the co-movements of couples' incomes (see Hyslop, 2001; and Shore, 2010, 2013). Apart from documenting that there is heterogeneity in the co-movement among couples, we also study how it evolves over the life-cycle and whether it differs over cohorts and skill groups. One important reason for why there is little research in this area is the data availability. The ideal dataset contains a long panel of individuals without problems of attrition or small sample sizes. For this study we use a dataset that can be said to approach this ideal, which is based on administrative data from Norway. The administrative data provides us with information on every Norwegian since several decades and we are able to link individuals to each other by marriage and children and to other records in order to find information on their characteristics and socioeconomic status. For this study we use information between 1975 and 2010 and we estimate the correlation of income shocks for each couple over time by first removing predictable changes to the individual incomes and computing the correlation of residuals for each couple over time with a 10-periods rolling window approach.

Similar to Shore (2013) we find that the co-movement of couples' income is close to zero, and we also show that there is heterogeneity among the couples. Households where both spouses have the same education level or are working in the same industry have a higher correlation of income shocks than households who have different education levels or work in different industries. This is also what we would expect if assortative matching on education and industry is important for the co-movement of shocks. The findings in relation to cohort, age and education indicate that there are small differences in average correlations. However, some patterns can be distinguished. Judging from the cohort plots it seems as if the co-movement has decreased over time, as earlier cohorts display a lower average correlation than more recent cohorts. Regarding the patterns over age, our results are somewhat in line with those of Shore (2013) who documents an increasing correlation of spousal income over the time in marriage. We find that the correlation increases slightly with age, but that it might be decreasing after a certain age for some groups. In particular, the increase over age is more pronounced for couples with lower education and the inverse u-shape is more pronounced for couples with high education. Together these patterns indicate that the insurance value of marriage may vary with time, age and education.

In a second step we compare these results to the evolution of factors that might explain part of the observed correlation patterns. In our dataset we have information on three factors that are expected to be important. These are labor supply as well as assortative matching on education and industry. We start with the assortative matching on education by studying plots of the share of households in which the individuals have the same education over cohorts

and education. We find that the assortative matching on education tends to increase over cohorts, i.e. younger individuals in our sample have a higher probability of having the same education as their spouse.[3] This might be a contributing factor to the increase in the income correlation that we observe over cohorts.

We would expect the assortative matching on industry to be a potentially stronger determinant of the co-movement of income shocks. In order to see whether patterns on assortative matching on industry may to some degree match the patterns of income correlation, we study how the likelihood of working in the same industry has evolved over cohorts, age and education levels. However, as we only have data on industry from 1993, we cannot study the entire evolution over age for different cohorts. The first interesting observation from this exercise is that the probability of working in the same industry is decreasing with cohorts. That is, individuals that are born later have a lower probability of working in the same industry. Thus, there are no descriptive evidence that assortative matching on industry can explain the increasing income correlation that we observe over cohorts. In relation to these results, however, it may be that the structure of industries has changed over time due to, for example, changes in technology and exposure to trade. Given the male or female dominance in certain industries it is not unlikely that these changes could have evolved differently for men and women, changing also the joint income dynamics of couples for different sectors over cohorts. When it comes to patterns over age and education there are some indications that assortative matching on industry may provide some explanation for the observed correlation patterns.

The final factor that we investigate is the importance of labor supply for the income correlation. Shore (2013) suggests that one mechanism behind the increasing correlation over time in marriage may be different patterns of substitution in production of spouses over age. In particular, specialization may be more important when there are small children in the household, but then the spousal time becomes more complementary over time until retirement. This would imply that the spousal labor supply correlation should increase over time. To study this we use the contractual annual hours of the spouses as a measure of labor supply. Unfortunately this measure is only available since 1999 and only for a part of the sample. Nonetheless, using these limited data we do observe that the correlation of spousal labor supply is increasing over age, potentially indicating increasing complementarity in production. In order to see how labor supply may impact the correlation of income shocks that we observe, we compare the correlation of wage shocks to the correlation of

[3]That the degree of assortative matching on education has increased over time has been found by, for example, Greenwood, Guner, Kocharkov and Santos (2014) for the US.

income shocks for the part of the sample for which we have labor supply data. Interestingly, this comparison shows that the correlation of wage shocks over age is steeper than the income correlation over age. At young ages the income correlation is higher than the wage correlation, whereas we observe the opposite case for older ages. This pattern is in line with an interpretation of gender-based productivity differences in home production, especially when the children are small. As such we could observe a lower correlation of labor supply at younger ages together with a higher income correlation and a lower wage correlation. It should however be emphasized that no firm conclusions can be drawn from this analysis due to the data limitations.

Our study is closely related to the scarce literature investigating the earnings dynamics of households. Hyslop (2001) seeks to understand how much of the increase in US inequality in the early 1980s that may be due to correlation in shocks to couples earnings and labor supply. Shore (2010) studies the intra-household risk-sharing over the business cycle. His results indicate that the covariance of couples' shocks to income is pro-cyclical and that marriage helps counter the increase individual risk in recessions. Shore (2013) studies how innovations to spouses' incomes change over time in marriage. His findings indicate that husbands with more volatile income tend to have wives with more volatile income and, similarly to our findings, that couples' income changes have a low correlation early in marriage and becomes more positive over time.

The current paper contributes to this literature by studying the association between spousal earnings using a long panel of data over individual earnings. Since we are using income data from the Norwegian tax registry, it is much less prone to measurement error than survey data, and this together with the long panel provides the possibility to estimate the correlation of spousal earnings with a higher degree of certainty. Moreover, with the current data we are not only able to study to what extent the income shock correlation changes over age and education, but also how it has evolved over cohorts. In addition, we look at how factors such as assortative matching on education and industry may affect these patterns.

This work is likely to have important policy implications. If the ability of households to provide income insurance varies in the population, for example among education and income groups, the appropriate policy responses to shocks in the economy may look different depending on where in the economy the shock hits.[4] Furthermore, that households have different insurance abilities because of differently correlated income shocks is important to

[4]In Chapter 3 of this thesis we confirm that spousal labor supply is an important insurance mechanism for households.

take into account when evaluating the effect of policies such as progressive income taxation and unemployment insurance.

The paper evolves as follows. In the next section we will present the data and in Section 2.3 we will go through the estimation approach. Section 2.4 presents the descriptive analysis and Section 2.5 concludes.

2.2 Data

2.2.1 Data description

The data that we use is obtained from several registry databases, which cover every individual in Norway and which we can link together using a unique individual identifier.[5] Using these databases we are able to contruct a dataset with the income of individuals for every year since 1975 and to link individuals to couples using information on who are married or who are cohabiting with children.[6] We can observe the latter since 1991. Through the databases we are also able to obtain information on individual characteristics such as age, education, household size, country of origin, county, and industry of the individual. We observe these variables from 1993. In addition we also have a measure of labor supply, namely the contractual hours of the individual. This variable shows the hours stated in the contract and thus it does not provide information on overtime or other hourly agreements not stated in the contract. We observe contractual hours for part of our sample since 1999.

In order to estimate the variance and the correlation of job related income risk in a couple we use the labor income of each individual. Labor income includes wage income and work-related transfers such as unemployment and sickness benefits, and maternity leave payments. The disposable income of the individuals, i.e. the income net of taxes, is calculated for each household using the relevant tax rates for each year. The income data is deflated using "grunnbeløp i folketrygden", which is a measure accounting for the inflation used to calculate for example social benefits. The income is expressed in Norwegian kroner (NOK) throughout the document. At this time 1 USD is around 6.6 NOK.

Our initial sample consists of the population of couples (as defined above) in Norway. These are 1,608,723 households and 27,141,660 observations, where one observation is household h in year t. Further we only consider individuals between the age of 25 and 60 to ensure that they have completed much of their schooling and too young to be eligible for early

[5]For more information on the Norwegian administrative data in general, see Røed and Raaum (2003).
[6]Unfortunately it is not possible to identify unmarried but cohabiting couples without children.

retirement. This restriction reduces the sample by 5,399,287 observations. From these observations we remove couples for which we are missing information on county, origin, or education (642,242 observations). As the focus of the paper is on the correlation of spousal income shocks, we restrict the sample to couples where both have positive labor market income and therefore drop 3,496,828 observations. We further require each household to have at least 15 years of positive income observations, which need not be consecutive (drop 5,541,928 observations).[7] The couples are allowed to break up and form new households which will be included if they conform with the other sample restrictions. The final sample consists of 546,213 households and 12,061,379 observations.

2.2.2 Descriptive statistics

Descriptive statistics for the sample are displayed in Table 2.1. The average household (disposable) income in the sample is around 524,000 NOK. The average male income at 326,000 is more than 1.6 times higher than the average income of females, which is 197,000 NOK. The average age of males is around 44 years old and of females around 41. The average household size in our sample is 3.7 individuals. On average females have fewer contractual hours per year than males at 1478 compared to 1834 for males. Note that since we only observe the contractual hours since 1999, the statistics shown are based on 1,683,870 couple-year observations while total sample contains 12,061,379 such observations. The average number of periods for which we observe a household is 23.5 years with a standard deviation of 5.4 years.

Before we go on to estimate correlation and variance of income shocks we will present the distribution of individuals in our sample over education and industry categories. The main purpose of this exercise is to get an idea of the extent of assortative matching on these variables in our data set. Assortative matching, in particular on industry, may be important for the extent to which income shocks are correlated within a household. The sample used for this exercise is the sample used in the regression analysis in Section 2.3.

Regarding education, the individuals are divided into fairly disaggregated education groups: less than secondary education; secondary education; some high school, high school degree; university lower/vocational degree or college education; university; and research. From Table 2.2 we can see how the individuals of the sample are distributed over education levels. Most males have reached a high school degree, whereas most females have a degree from a non-scientific education such as teaching or nursing, or another professional degree.

[7] Using only 10 years does not change the conclusions of the paper, but it adds more noise.

Table 2.1: Summary statistics

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Household income	523795.0	235469.6	2.4	198267168
Income, female	197377.2	103847.9	1.1	15540156
Income, male	326417.8	193110.8	1.1	197877072
Age, female	41.5	8.5	25	60
Age, male	43.7	8.6	25	60
Household size	3.7	1.1	2	6
Hours of work, female	1477.8	524.7	195	2353
Hours of work, male	1834.4	337.1	195	2353
No of years per household	23.5	5.4	15	36
N		12,061,379		

Note that the labor supply statistics are based on 1,683,870 observations.

To get an idea of the degree of assortative matching in our sample we study how many individuals who have the same education level as their partner. On average this share is around 36 percent of the couples in our sample. Moreover, column four of Table 2.2 shows the probability of male and female having the same education level conditional on the education of the male.

Table 2.2: Distribution over education

Education	$educ_m$	$educ_f$	$P(educ_f = educ_m \mid educ_m = j)$	$P(educ_m = j \mid educ_f = educ_m)$
<Secondary	0.1	0.2	22.6	0.1
Secondary	16.4	18.4	35.4	16.3
Some High School	21.6	27.9	41.7	25.2
High School	26.4	18.4	25.0	18.5
Uni lower/Vocational degree	24.5	30.3	49.3	33.9
University	9.7	4.2	20.7	5.7
Research	1.2	0.5	11.6	0.4
Sum	100	100		100

This probability is higher than 20 percent for all education categories apart from research where it is almost 12 percent. The probability is highest for the case when the male has a professional degree (49.3 percent). The fifth column of Table 2.2 instead shows the probability of the male having a specific education level conditional on the individuals having the same level of education. This probability ranges between 0.4 for research and around 34 percent for some college education. From this we can conclude that there is a considerable degree of assortative matching when it comes to education levels.

In order to do the same analysis for industries we divide individuals into industries using

the 2-digit NACE code (revision 2), which we observe for our sample since 1993. It is not shown here, but at this level of aggregation, around 17 percent of the households work in the same industry. From Table 2.3 we observe that most males work in manufacturing (19.5 percent), followed by wholesale and retail trade (13.9 percent) and construction (10.9 percent). Water supply, sewerage and waste management is an industry with few workers in general, and 0.7 percent of the males works in this industry. For the females in our sample the industry choice seems to be more concentrated. The far most popular industry is human health and social work with 33.2 percent of the females. This category is followed by education (13 percent) and whole and retail trade (12.7 percent). As for the males, not many females work in water supply, sewerage and waste management (0.1 percent).

Furthermore, column four shows the probability that the individuals in a couple work in the same industry conditional on the male industry. This probability ranges between 3.5 percent for water supply, sewerage and waste management and 54.6 percent for human health and social work. If the male works in education the probability of working in the same industry is 36.6 percent. Accommodation and food services has a low share of both males and females, but a high probability that both spouses work there at 31.7 percent, the opposite case from water supply, sewerage and waste management.

Column five shows the probability that the male works in a particular industry given that both spouses work in the same industry. This probability is highest for wholesale and retail trade (16.2) and lowest for water supply, sewerage and waste management (0.1 percent).

2.3 Estimation of income variance and correlation

In the following we will estimate the couple income correlation and variance.

As mentioned above, we have at least 15 non-zero observations of labor income for each individual in the household. We use the series of individual disposable incomes in a couple to estimate the total variance and the correlation of household income shocks. In a first step we purge the individual labor income of predictable factors. In order to do this we regress the natural logarithm of individual income on age and age squared, dummies for the education level, origin, county, and year, as well as the interaction of education with age and age squared and education and time.

The regression model:

$$\ln y_{hit} = \beta_0 + \mathbf{X}_{hit}\beta_i + county_{ht} + \tau_t + \epsilon_{hit} \tag{2.1}$$

Table 2.3: Distribution over industries

Industry	ind_m	ind_f	$P(ind_f = ind_m \mid ind_m = j)$	$P(ind_m = j \mid ind_f = ind_m)$
Agriculture, forestry and fishing	1.5	0.6	10.0	0.8
Mining and quarrying	3.0	0.7	6.6	1.1
Manufacturing	19.5	7.2	15.2	16.2
Electricity and gas	1.7	0.4	3.7	0.3
Water supply, Sewerage and waste management	0.7	0.1	3.5	0.1
Construction	10.9	1.5	6.0	3.6
Wholesale and retail trade	13.9	12.7	24.9	18.9
Transportation and storage	8.7	3.3	10.0	4.8
Accommodation and food service activities	1.0	2.1	31.7	1.7
Information and communication	1.2	0.6	5.8	0.4
Financial and insurance activities	2.9	3.1	12.0	1.9
Real estate activities	1.1	0.7	8.6	0.5
Professional, scientific and technical activities	3.0	1.4	8.6	1.4
Administrative and support service activities	6.4	5.5	14.7	5.1
Public administration and defense	9.3	10.7	21.1	10.8
Education	7.1	13.0	36.6	14.2
Human health and social work	5.5	33.2	54.6	16.5
Arts, entertainment and recreation	1.2	1.1	11.0	0.7
Other service activities	1.5	2.0	11.6	1.0
Sum	100	100		100

$\mathbf{X}_{hit}$ includes the observable characteristics of individual i in household h at time t, where i represents either the female or the male. τ_t is a year dummy, and ϵ_{hit} is the error term. The model is estimated by OLS and from these regressions we obtain the residuals ($\hat{\epsilon}_{hft}$ and $\hat{\epsilon}_{hmt}$).

For the estimation of the variance of total household income we use the residuals from a regression of the natural logarithm of total household labor income (the sum of male and female income) on the individual characteristics of *both* the male and the female included in the individual regressions as well as the dummies for county and time.[8]

$$\ln(y_{hft} + y_{hmt}) = \beta_0 + \mathbf{X}_{hft}\beta_f + \mathbf{X}_{hmt}\beta_m + county_{ht} + \tau_t + \epsilon_{ht} \tag{2.2}$$

It is likely that the variance and correlation vary over cohort, age and education.[9] We would like to capture this in our estimates of the variance and the correlation, using the household as a basis for estimation. Therefore, rather than producing one point estimate of

[8]We are not controlling for household size here as we only have data on this variable since 1993.

[9]For example, Findings by Blundell, Graber and Mogstad (2014) show that the variance of male income vary considerably over age, time and education.

the variance and the correlation for each household we estimate the correlation and variance using a rolling window approach in order to obtain several estimates of variance and correlation for each household. More specifically, for each household we estimate the variance and correlation in year, t, using the income information between $(t-8:t+1)$. In order for the household to have a non-missing variance and correlation in any given year, t, the household needs a minimum of 5 years where both spouses have a positive income within the period $(t-8:t+1)$. Since we observe each household in the dataset for at least 15 years, one household should have at least 6 estimates for the variance and the correlation if none of the estimates are missing.[10]

The estimation results are shown in Table 3.2.[11] The variance of female income is greater than the variance of male income at 0.218 compared to 0.106 and the female variance has a larger standard deviation than that of male income. Similar to the findings of Shore (2013), our estimation results indicate that the co-movement of shocks is close to zero with an average correlation of 0.055. The standard deviation of the correlation is 0.435 and the range stretches between -0.896 and 0.9. The variance of total household income is estimated to be 0.033 with a standard deviation of 0.099.

Table 2.4: Summary statistics variance and correlation of income shocks

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Variance female income	0.218	0.560	0	35.78
Variance male income	0.106	0.398	0	32.681
Correlation male and female income	0.055	0.435	-0.896	0.9
Variance household income	0.033	0.099	0	15.26
N		9,274,853		

2.4 Co-movements of labor income shocks

In this section we provide a descriptive analysis of the patterns of correlation as well as relate the patterns to some of the factors that might influence the co-movement of labor income shocks.

[10]For now we do not attempt to distinguish permanent from temporary shocks to income as it requires more information. Instead we are focusing on estimating time varying measures of the variance and the correlation for each household.

[11]Note that the number of household-year observations is lower for these summary statistics than for the ones shown above because the estimations of the variances and the correlation are only from 1983. The years before will have missing values.

2.4.1 Correlation over education and industry

If assortative matching on education or industry matters for the co-movement of shocks to the spousal incomes we would expect that spouses who have the same education or who work in the same industry would on average have a higher correlation, and perhaps variance, of these shocks than spouses who do not. To study this in our sample we compare averages of the income variance and correlation of couples who have the same education level to couples who have different education levels, as well as averages of income variance and correlation of couples who work in the same industry to couples who work in different industries. The education and industry groups are represented by the male education level and industry. As mentioned before, we only observe industry since 1993 and thus, part of our total sample is not included in the comparison of industries.

We start with education. The first two columns of Table 2.5 show the mean correlation and the two latter columns display the mean variance. Furthermore, column 1 and 3 show averages by education level for couples who have *different* education levels and column 2 and 4 show the average estimates for the households where the spouses have the *same* education level. Note that subscript h denotes household h.

Studying Table 2.5 we see that the average correlation displays a u-shape over the education levels; it is highest for couples with a low and a high level of education and lowest for high school. If we compare couples with a different education level (column 1) to couples with the same education level (column 2) we notice that the average correlation is higher for couples with the same education in all cases, but for high school. All differences are significant at the 1 percent level, except for the couples with less than secondary education where the difference is significant at the 10 percent level, and for individuals with a high school degree in which case we cannot reject the hypothesis that they have the same average correlation.[12]

Regarding the total income variance of the household, it more or less follows the u-shaped pattern of the correlation. However, we note from the last row of Table 2.5 that the couples with the same education actually have a lower total variance than couples with different education on average, although we cannot reject the hypothesis that the average variance is the same. For the following education groups, the total variance is higher for those with the same education: less than secondary; secondary; and university. For the other groups we see the opposite result, but for research we cannot reject the hypothesis that the groups have the same average. This means that even if the couples with the same education tend to have

[12]All the test results reported in this section are not displayed in the paper, but can be obtained by the authors upon request.

a higher correlation, it seems to some extent offset by lower variance of the individual labor incomes.

Table 2.5: Averages of variance and correlation of income shocks over male education levels

	$Mean(corr(\hat{\epsilon}_{ft}, \hat{\epsilon}_{mt}))$		$Mean(var(\hat{\epsilon}_{ht}))$	
	$educ_f \neq educ_m = j$	$educ_f = educ_m = j$	$educ_f \neq educ_m = j$	$educ_f = educ_m = j$
< Secondary	0.0290	0.0523	0.0419	0.0770
Secondary	-0.0144	0.0003	0.0251	0.0282
Some High school	-0.0046	0.0162	0.0244	0.0222
High school	-0.0054	-0.0063	0.0219	0.0211
Uni lower/vocational degree	-0.0044	0.0006	0.0241	0.0225
University	0.0099	0.0738	0.0248	0.0358
Research	0.0533	0.1180	0.0294	0.0291
Total	-0.0036	0.0069	0.0238	0.0237

Note: Column (1) and (3) refer to households which have different education level, whereas column (2) and (4) refers to households with the same education level.

Table 2.6 displays the averages over industry groups. Column 1 and 2 display the correlation and column 3 and 4 show the variance. In column 1 and 3 we see the averages by industries for couples who work in different industries and in column 2 and 4 we find the average estimates for couples who work in the same industry. When studying the table it is important to keep in mind that the variance displayed is not the variance of male income for each industry, but of the total household income for each male industry (residual denoted by h). Nonetheless, from column 1 we can see that in case the male works in some of the industries considered risky, such as agriculture and real estate activities, the total income variance is relatively high. If the male instead works in industries associated with the public sector, such as public administration and defense, as well as education, the total variance tends to be lower (0.017 and 0.019 respectively). From Table 2.3 we know that for the two latter industries, the probability of the female working in the same industry conditional on the male working there is quite high. Therefore, the averages of total income displayed in the table can be expected to be low due to low variances in both individual incomes.

Comparing the average variance of couples working in different industries to couples working in the same industry (column 3 and 4 Table 2.6) we note from the last row of the table that, on average, the couples working in the same industry have a higher total variance and that the difference is significant at the 5 percent level. Furthermore, for 12 out of the 18 industries for which a comparison is possible, couples who work in the same industry have a higher variance on average, and the difference is significant at least at the 5 percent level. The exception is accommodation and food service activities where we cannot reject the hypothesis that the average variance is the same between the groups. In the cases

where the couples working in the same industry have a lower variance (electricity and gas, information and communication, financial and insurance activities, public administration and defense, education, and human health and social work), all differences are significant at the 0.1 percent level, except for electricity and gas where the difference is significant at the 10 percent level. It is interesting to note that half of these industries are industries associated to the public sector where we would expect a low variance of income shocks.

When it comes to the correlation of couples working in the same industry, we see from column 2 that the highest correlation can be found in agriculture, forestry and fishing (0.201), followed by accommodation and food service activities (0.163). The lowest correlation for couples working in the same industry can be found in transportation and storage (0.024) and in human health and social work (0.031). Comparing column 1 to column 2 we see that, as expected, the average income correlation of couples working in the same industry is higher than the average income correlation of couples working in different industries for all industry groups. All differences are significant at the 0.1 percent level. The only industry where a comparison is not possible is for 'water, sewerage and waste management' where we, due to few observations, did not obtain an estimate for couples working in the same industry.

Table 2.6: Averages of variance and correlation of income shocks over male industry

	$Mean(corr(\hat{\epsilon}_{ft}, \hat{\epsilon}_{mt}))$		$Mean(var(\hat{\epsilon}_{ht}))$	
	$ind_f \neq ind_m = j$	$ind_f = ind_m = j$	$ind_f \neq ind_m = j$	$ind_f = ind_m = j$
Agriculture, forestry and fishing	0.0319	0.2010	0.0392	0.0723
Mining and quarrying	-0.0450	0.0372	0.0203	0.0224
Manufacturing	-0.0196	0.0384	0.0196	0.0220
Electricity and gas	-0.0295	0.0376	0.0129	0.0118
Water supply, sewerage and waste management	-0.0077	-	0.0201	-
Construction	-0.0072	0.0954	0.0242	0.0451
Wholesale and retail trade	-0.0177	0.0502	0.0230	0.0341
Transportation and storage	-0.0128	0.0240	0.0247	0.0251
Accomodation and food service activities	0.0112	0.1630	0.0355	0.0632
Information and communication	-0.0178	0.0366	0.0174	0.0147
Financial and insurance activities	-0.0191	0.0458	0.0161	0.0140
Real estate activities	-0.0069	0.0981	0.0333	0.0675
Professional, scientific and technical activities	-0.0053	0.0994	0.0238	0.0302
Administrative and support service activities	-0.0153	0.0748	0.0312	0.0473
Public administration and defense	0.0058	0.0332	0.0188	0.0172
Education	0.0082	0.0671	0.0224	0.0192
Human health and social work	0.0065	0.0312	0.0298	0.0286
Arts, entertainment and recreation	-0.0044	0.0608	0.0267	0.0361
Other service activities	-0.0150	0.0989	0.0262	0.0419
Total	-0.0112	0.0508	0.0229	0.0278

Note: Column (1) and (3) refer to households which have different industry classifications, whereas column (2) and (4) refers to households with the same industry classification.

From this analysis we conclude that working in the same industry and, to some extent, having obtained the same educational degree, is important for the co-movement of couples' income. This indicates that assortative matching on these variables could also be important for the income risk that a household faces as the correlation of income shocks is likely to tell us something about the insurance ability of a household.

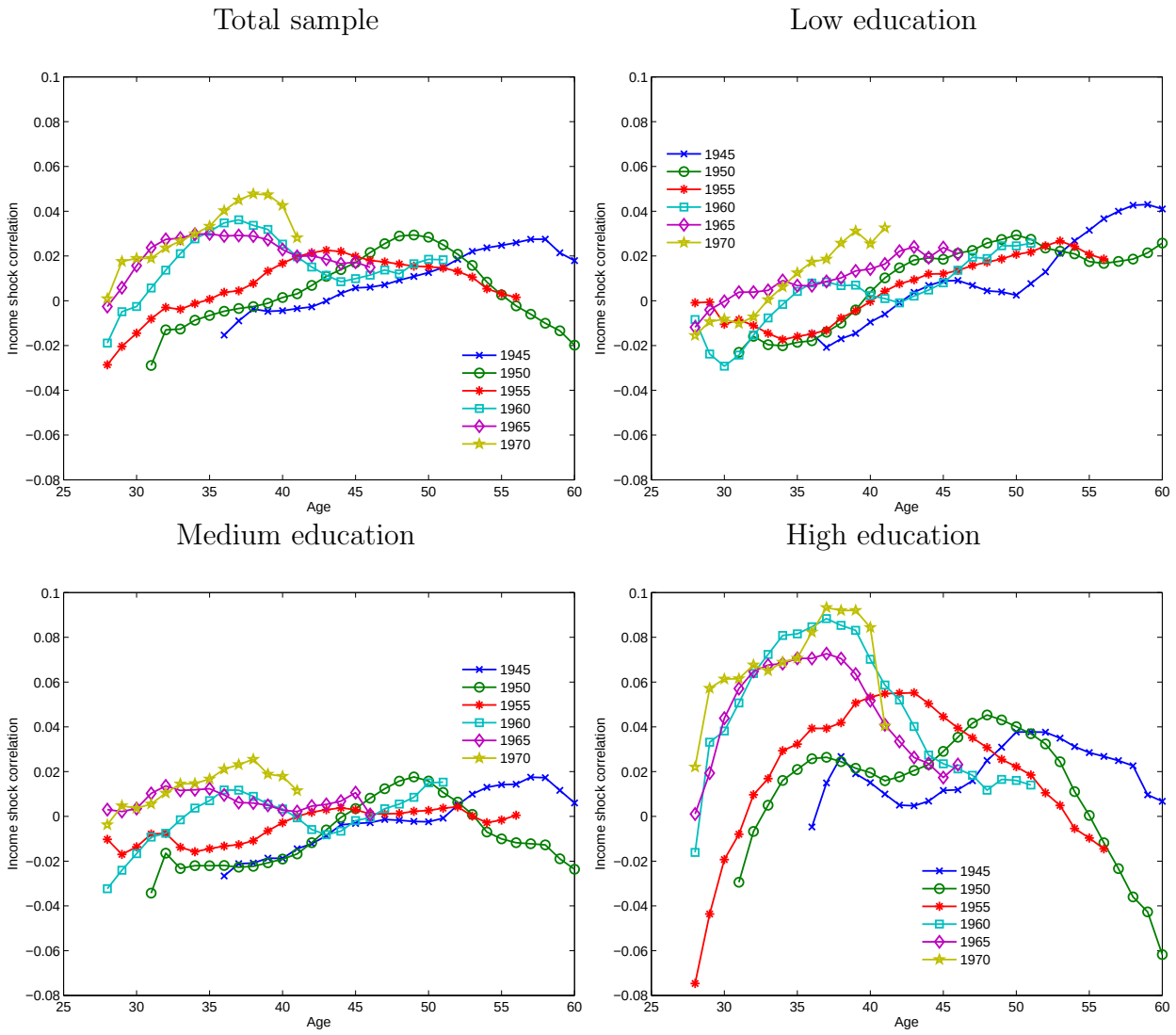
2.4.2 Income correlation over cohort and age

Since we have data of the income of individuals over time since 1975 we can study potential differences in income shock correlation over cohort and age. Figure 2.1 shows these patterns for the total sample and for different education levels. In the upper left panel the development of the correlation for the total sample is pictured. Overall, the correlation does not seem to change much over the life time or between cohorts, but some patterns can be distinguished. First of all, the correlation tends to increase over age at earlier ages until it reaches a turnings point and starts a slight decrease. These findings are similar to those of Shore (2013) who finds that the correlation of couples' incomes tend to increase over time in marriage. There are some differences between cohorts in that more recent cohorts have a slightly higher correlation and the point at which the correlation starts decreasing moves with cohorts towards earlier ages.

For the other panels in Figure 2.1 we have divided the couples into low, medium, and high education. The division is based on the male education. Low education includes secondary education or less, medium education comprises all individuals who went to and or graduated from high-school, and high education are individuals who have gone to college, or graduated from college, as well as individuals who went on to research. There seems to be differences in correlation patterns for the different education groups. For the low educated the correlation is more clearly increasing over age, while the correlation seems more constant over time for the medium educated and we can distinguish a slight hump-shaped pattern. In the lower right panel we see the same figure for the highly educated and their pattern is remarkably different from the other education levels. First of all, there is a greater variance in the correlation. Secondly, apart from the earliest cohort, for which the correlation is relatively constant, the correlation displays a clear hump-shape over age. The three later cohorts have a higher correlation for the earlier ages, but the correlation seems to decrease rapidly after the late 30s. One explanation for the strong increase in correlation maybe that the individuals get established on the labor market later and they might have children later, contributing to a decrease in correlation a bit later. Similarly, that lower educated in general have children at earlier ages may explain why their correlation is more clearly increasing. In

general the hump-shape may have something to do with issues of sample composition. Hess (2004) finds that a higher income correlation tends to increase the probability of divorce. If this is the case also in our sample we should see a decrease in the correlation over age as the remaining couples have a lower correlation.

Figure 2.1: Correlation over cohort and age



2.4.3 Potential determinants of income correlation over cohort and age

From Figure 2.1 we may be able to distinguish different patterns in the correlation of income shocks over cohort, age and education. In this subsection we study some of the factors that might influence this correlation. Likely candidates are assortative matching on education and on industry, as well as labor supply.

Assortative matching on education. As shown in table 2.5, having the same level of education may increase the probability of a higher correlation. Before we study the patterns of the share in same education more closely we will note one thing in relation to sample composition. Studying the correlation patterns in Figure 2.1, we may be worried that changing sample composition could explain part of the variation in correlation. In general we expect to have fewer observations for younger ages when individuals are obtaining education and at older ages when some couples start to retire. If the couples who enter later or exit earlier have characteristics that differs in ways relevant to the correlation, this would be captured in the figures.[13] If we believe that assortative matching on education may capture such characteristics, we could use Figure 2.2 to study this issue. Figure 2.2 shows the share of couples with the same education over age and cohort for different male education levels. Since we use the highest obtained education as our education measure and as individuals in our sample are not students, we would expect the share of individuals in the same education for each cohort to be constant over age if sample entry and exit is unrelated to the probability of having the same education. From Figure 2.2 we can see that the share of couples with the same education is fairly constant over age when looking across cohort and education. One exception is the highly educated, where the share in the same education of the two most recent cohorts actually follow the same hump-shape as the income shock correlation correlation. Otherwise, it does not seem to be much important sample composition changes according to this measure.

In relation to the correlation patterns we may note that there are significant differences in assortative matching on education between the cohorts and the education levels. Couples where the male has medium education are more likely to have the same education and couples where the male has high education are the least likely. Furthermore, there are interesting differences between cohorts. For both medium and highly educated there is a clear increase in assortative matching on education with more recent cohorts showing a higher probability

[13]For example, Hess (2004) shows that a marriage between two individuals who have a higher correlation in income and a larger difference in income volatility is more likely to end earlier.

of having the same education. The households with a medium educated male show an average increase from 0.3 for the 1945 cohort to around 0.5 for the 1970 cohort, and for the households with a highly educated male the same numbers are 0.15 and 0.3. These patterns may contribute to the increase in co-movement of income risk over cohorts shown in Figure 2.1. On the other hand, the pattern for the least educated is the exact opposite. The share in same education of the 1945 cohort is around 0.45 and for the 1970 cohort 0.2. To some extent, the pattern of the share in the same education possibly reflects the increase in education attainment of females. Figure 2.3 shows a slight increase of assortative matching on education over cohorts overall. This may contribute to an explanation of why more recent cohorts have a higher income shock correlation in our sample.

Figure 2.2: Share in same education over cohort and age

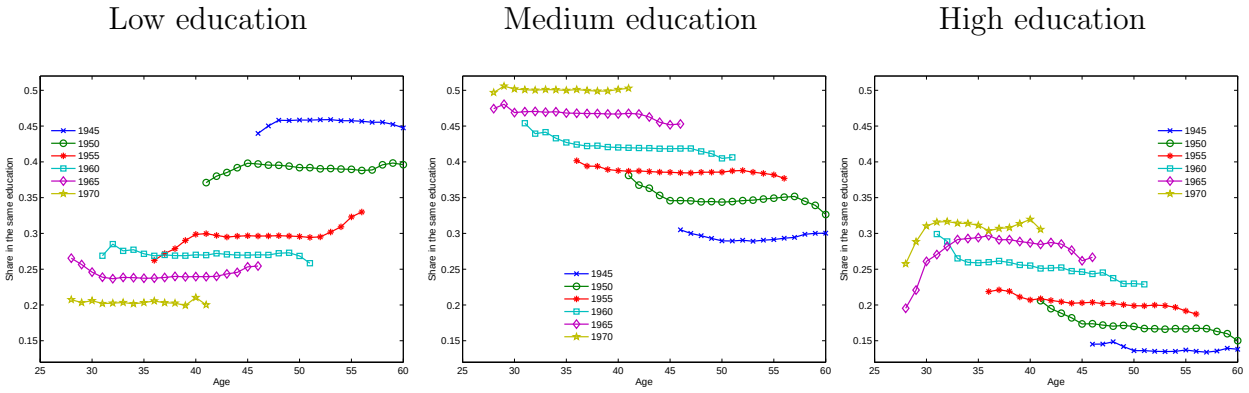
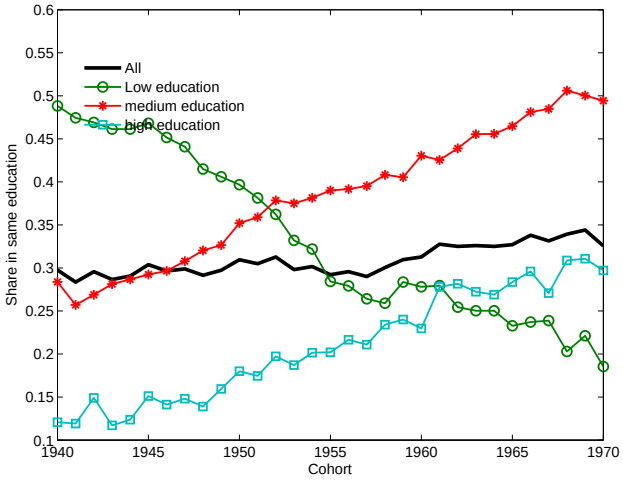


Figure 2.3: Share with same education over cohorts



Assortative matching on industry. As we saw in Table 2.6, individuals who work in the same industry have a higher income shock correlation on average than individuals who do not. In order to see to what extent assortative matching on industry may be reflected in the correlation patterns of Figure 2.1, we plot the share of individuals in the same industry in the same way. These patterns are pictured in Figure 2.4. In general, we do not expect the industry of an individual to be the same throughout life and we can see that the share of couples in the same industry is varying slightly over age. In the upper left panel we picture the total sample from 1993 (since we only have industry information from this point). Moreover, we may note that the share of couples working in the same industry almost follows the same hump-shape over ages as the correlation in Figure 2.1. The hump-shaped pattern could be due to changing employment opportunities in the economy that affects all cohorts in the same way at the same time.[14] Differently from what we would expect from the pattern of correlation, earlier cohorts have a higher tendency to work in the same industry. This may partly be due to the increase in the number of different industries over time. Moreover, the share seems to be more constant for earlier cohorts, whereas it is slightly decreasing for younger cohorts over earlier ages. With the number of industries fixed, this may to some extent reflect a greater tendency of later generations to change jobs over the life time, which might imply changing industry as well. Turning to differences over education levels (the other three panels), we may note that the probability of working in the same industry seems to increase in the male education level, which may contribute to the slightly higher correlation of couples with a higher male education noted in Figure 2.1. Moreover, especially for couples where the male has a low education level, the share working in the same industry tends to decrease with age, implying that there are other factors than industry that might be important for the correlation pattern of low educated. The low educated also have a less pronounced hump-shape in the share with the same industry than the other education levels, which may be reflected in Figure 2.1 as their income shock correlation displays less of an hump-shape. For couples where the male has a high education the two earliest cohorts show an increase in the share working in the same industry towards older ages. One interpretation from the education figures could be that the households to some extent try to increase the risk-sharing potential of the marriage over time by changing industry to one where the partner does not work. To the extent that couples with a higher education have more secure

[14] Another candidate explanation is revisions to the NACE codes that have taken place over time, mostly to reflect an increasing number of industries. However, these changes took place in 2002 when NACE rev. 1 became NACE rev. 1.1 and in 2006 when it became NACE rev. 2 (see Eurostat, 2008). These are not the years when we observe the hump-shape. We have also worked in making the codes as consistent as possible over the sample.

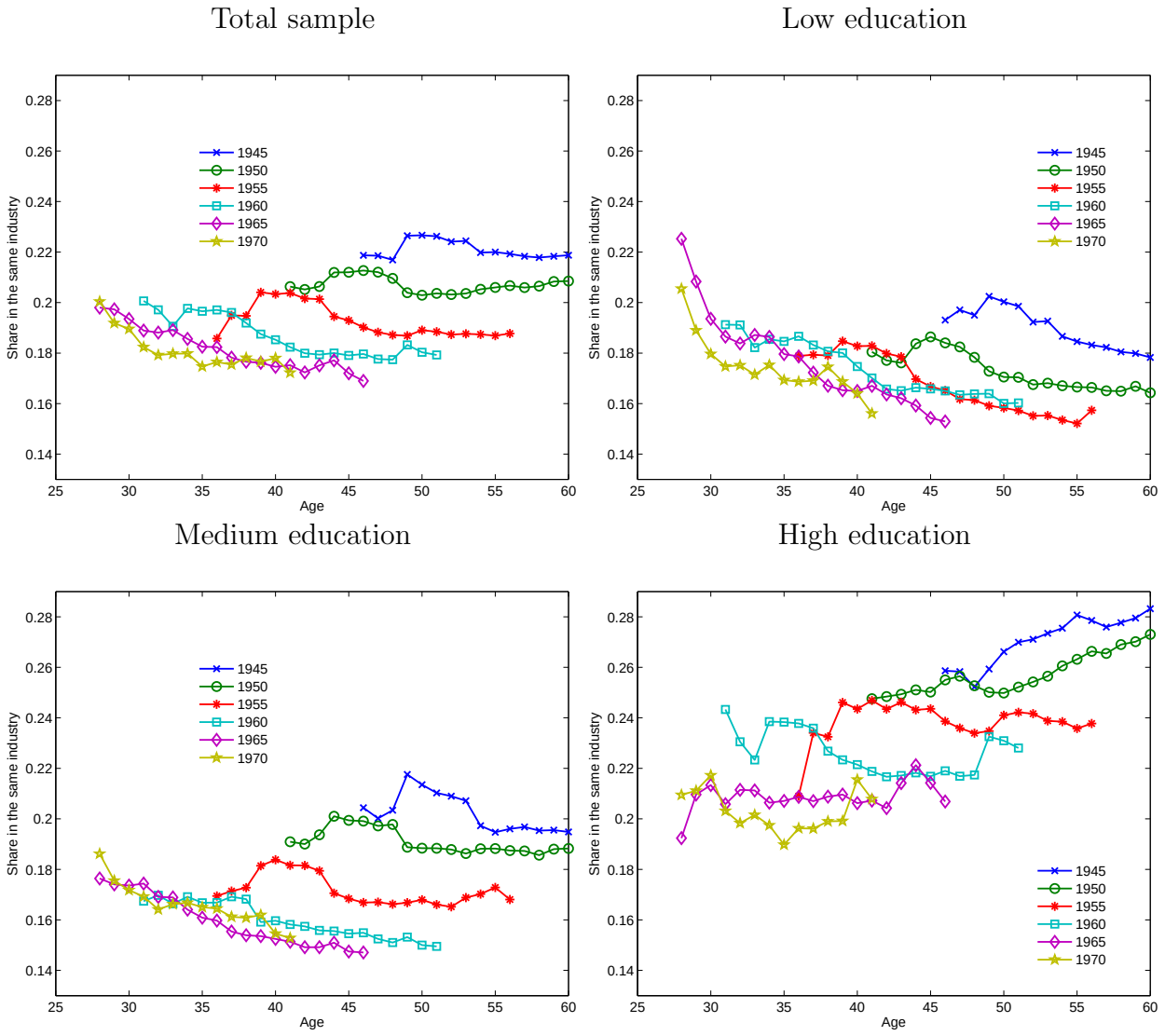
jobs, better alternatives in case of job loss, or larger buffers in general, this action may be more important for lower educated individuals.

Although Figure 2.4 indicates that working in the same industry may not be an important factor in explaining the differences in correlation over cohorts, it might have something to say in relation to the pattern of correlation over age and education. We cannot see the whole picture, however, as we lack industry information for the younger ages of the earlier cohorts. Moreover, in regards to cohort patterns, apart from an increasing number of industries over time, one could imagine that the characteristics of industries have changed over time due to, for example, changes in technology or trade exposure. Such changes may impact the strength of the association between working in the same industry and the income shock correlation. For males and females in the more recent cohorts, for example, working in the same industry may imply having more similar occupations than it did for earlier generations, and therefore the older generations could have a lower correlation even if they work in the same industry. Unfortunately we do not have information on occupation so we cannot check this hypothesis for our sample.[15]

Labor supply. As mentioned, one potential reason for the slight increase in income shock correlation over age may be, as pointed out by Shore (2013), that spouses specialize in production when there are young children in the household and that their time becomes more complementary as they approach retirement. For the main part of the sample, the theory would predict an increasing correlation in labor supply over marriage. As mentioned in the data section, our dataset does not contain an exact measure of labor supply, but from 1999 we know the contractual hours of part of the individuals in our sample. Figure 2.5 pictures the result of estimating the correlation of couples' labor supply over time, using the rolling windows procedure of before, but with 7 periods instead of 10. As we can see from the figure, the labor supply correlation of spouses indeed seems to increase over age and the pattern seems to be very similar for all education levels. There is a steeper increase in the labor supply correlation of highly educated over earlier ages which could explain part of the steep increase in income correlation of these couples. Thus, increasing complementarity of production may explain part of the slight increase we observe in the income correlation. However, what should be remembered is that we cannot see the whole development over age for each of the cohort due to the few time periods for which we have data. This becomes clear from the right panel of Figure 2.5 which shows the labor supply correlation divided

[15]In Chapter 3 we use the interaction of male industry, female industry and time as an instrument for the correlation in income shocks estimated at household level. This interaction has a very strong predictive power in the first stage.

Figure 2.4: Share working in the same industry over cohort and age



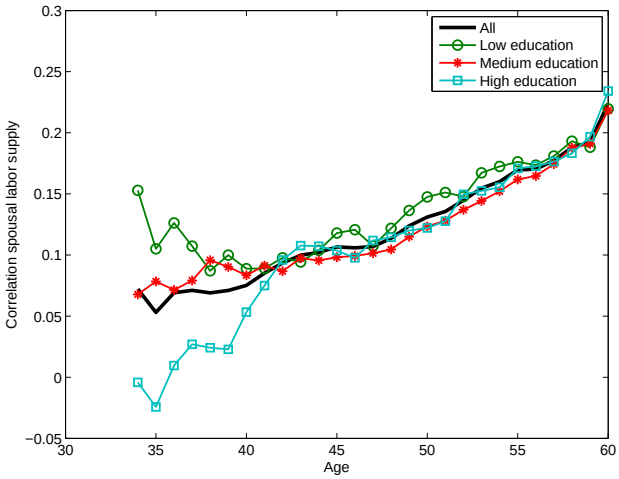
into cohorts.

In order to try to see how much the correlation pattern may change when taking labor supply into account, we use the labor supply information to estimate wage residuals, using the same procedure as specified in equation 3.1 and equation 3.2.

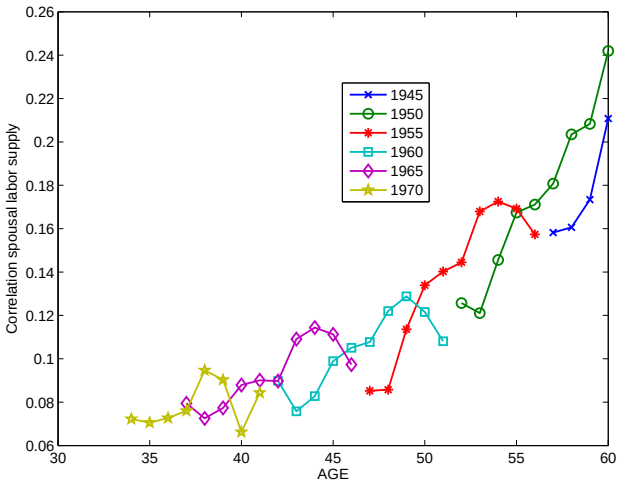
As we have a short time period available we estimate the income correlation with a rolling windows procedure using 7 years, as for the labor supply correlation above. Since the estimates are less precise using 7 periods in the estimation, we plot the curves using moving averages over three age categories in order to better distinguish patterns. The result for this part of the sample (with information since 1999) is shown in Figure 2.6. The first

Figure 2.5: Correlation spousal labor supply

Over age and education



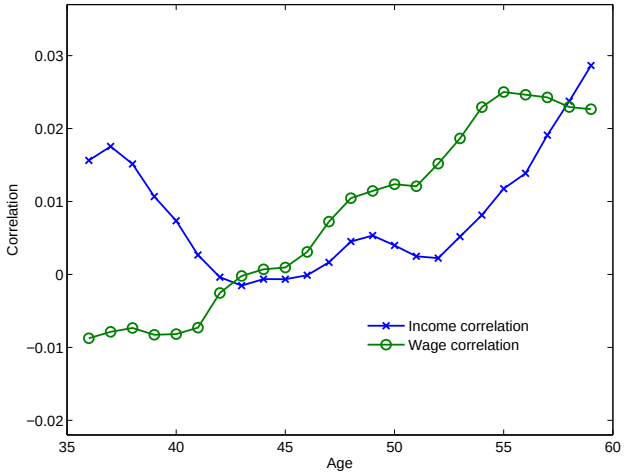
Over age and cohort



thing to note from the figure is that the pattern of correlation looks somewhat different as compared to before. This may be the result of several cohorts and education groups together and, furthermore, of sample composition issues as we do not have labor supply data for all households that we observe since 1999. Nevertheless, the figure may provide an indication of how labor supply changes the correlation pattern. Overall, the differences between income and wage correlation is not that big and, apart from at earlier ages, the curves follow each other. Moreover, there seems to be smaller variation in the income correlation than in the wage correlation and, interestingly, the wage correlation is steeper over ages than the income correlation. This means that the increase in correlation of spousal labor supply over age cannot explain the increase in the correlation of income shocks. One potential reason for why the correlation in wages and labor supply is low, while the correlation in income shocks is high at younger ages is gender-based productivity differences in household production, especially at the time when there are small children. If the female is much more productive in taking care of children and the household, this is likely to be reflected in labor supply patterns, potentially even for larger positive relative female wage shocks.

Even though we may not be able to draw strong conclusions regarding the effect of labor supply on the income correlation due to the scarcity of data, the results from this section point in interesting directions for future research.

Figure 2.6: Comparison income and wage correlation over age



2.5 Conclusion

As described in the introduction to this chapter, the co-movement in income shocks of a household is likely to provide an indication of the ability of the household to insure labor income through income pooling and labor supply. In this paper we provide evidence that there is heterogeneity in the population when it comes to co-movements in income shocks, measured by their correlation. As expected we find that couples working in the same industry and, to a lesser extent, couples with the same education, have a higher correlation than the other couples. Furthermore, according to our findings the correlation varies with age, cohort and education. More specifically we find that the correlation is slightly negative for young households around the age of 25, that it increases somewhat with age and, in some cases, decreases again. Households where the male has a low education experience a more pronounced increase in the correlation over age on average, and households where the male has a high education have a more pronounced hump-shape. Furthermore, the correlation tends to increase over cohorts with more recent cohorts showing a higher correlation. As such, the results indicate that the insurance ability of households may have changed over time and that it changes with age and with education levels.

When studying potential factors behind the correlation patterns, we find that the correlation seems to vary with cohort, age and education in a way that might be partly explained by assortative matching on education and industry. Moreover, for part of the sample we are able to compare income, wage and labor supply correlations. The labor supply correlation tends to increase with age, pointing to increasing complementarity in production with

the age of the spouses. The wage correlation displays a steeper increase over age than the income correlation, further indicating that there may be gender-based differences in home production, especially when children are small. However, as we only have scant data on labor supply we cannot look at patterns for the whole sample for a longer period of time.

One limitation of this analysis is that it is purely descriptive, which means that it is difficult to distinguish the actual influence of different factors on the correlation of income shocks. Moreover, the current analysis may be sensitive to entry and exit of households that may change the sample composition in ways relevant to the analysis. In future work we will attempt to see how much of the variation in correlation between different groups that can be accounted for by the different factors in a more rigorous way.

Bibliography

[1] Blundell, R., Graber, M. & Mogstad, M. (2014). Labor income dynamics and the insurance from taxes, transfers and the family. *Forthcoming in Journal of Public Economics.*

[2] Blundell, R., Pistaferri, L. & Saporta-Eksten, I. (2012). Consumption inequality and family labor supply. NBER Working Paper, No. 18445.

[3] Cullen, J.B. & Gruber, J. (2000). Does unemployment insurance crowd out spousal labor supply? *Journal of Labor Economics, 18*, 546-572.

[4] Eurostat (2008). NACE Rev. 2 - Statistical classification of economic activities in the European Community. Eurostat Methodologies and Working papers.

[5] Fagereng, A., Gottlieb, C. & Guiso, L. (2013). Asset market participation and portfolio choice over the life-cycle. Netspar Working Paper.

[6] Greenwood, J., Guner, N., Kocharkov, G. & Santos, C. (2014). Marry your like: Assortative matching and income inequality. NBER Working Paper, No. 19829.

[7] Gottschalk, P. Moffitt, R. (1994). The growth of earnings stability in the U.S. labor market. *Brookings Papers on Economic Activity, 112*, 217-272.

[8] Hess, G.D. (2004). What's love got to do with it? *Journal of Political Economy, 112*, 290-318.

[9] Hyslop, D.R. (2001). Earnings inequality and family labor supply. *American Economic Review, 91*, 755-777.

[10] Juhn, C. & Potter, S. (2007). Is there still an added worker effect? Staff Report, Federal Reserve Bank of New York. No. 310.

[11] Lundberg, S. (1985). The added worker effect. *Journal of Labor Economics, 3*, 11-37.

[12] Moffitt, R. Gottschalk, P. (1995). Trends in the autocovariance structure of earnings in the U.S., 1969-1987. Mimeo, Johns Hopkins University.

[13] Moffitt, R. Gottschalk, P. (2002). Trends in the transitory variance of earnings in the United States. *The Economic Journal, 112*, C68-C73.

[14] Moffitt, R. Gottschalk, P. (2006). Trends in the earnings volatility in the U.S.: 1970-2002. Working Paper.

[15] Ortigueira, S. & Siassi, N. (2013). How important is intra-household risk sharing for savings and labor supply? *Journal of Monetary Economics, 60*, 650-666.

[16] Roed, K. & Raum, O. (2003). Administrative registers - unexplored reservoirs of scientific knowledge? *The Economic Journal, 113*, F258-F281.

[17] Shore, S.H. (2010). For better, for worse: Intra-household risk-sharing over the business cycle. *Review of Economics and Statistics, 92*, 536-548.

[18] Shore, S.H. (2013). The co-movements of couples' incomes. *Review of Economics of the Household*, May 2013.

[19] Stephens Jr., M. (2002). Worker displacement and the added worker effect. *Journal of Labor Economics, 20*, 504-537.

Chapter 3

Household Risk, Labor Supply and Precautionary Wealth Accumulation

With Andreas Fagereng and Kjersti Naess Torstensen

3.1 Introduction

When it comes to income fluctuations and employment risk the individual relies to a large extent on labor supply and precautionary savings for insurance. The main reasons are that private insurance markets do not exist, that individuals are credit constrained, and that the insurance coverage provided by governments in case of loss of employment is limited.[1] Another potential source of income insurance is the marriage itself through income pooling and spousal labor supply. Lundberg (1985) investigated the insurance potential of the latter channel in what she called the "added worker effect", but spousal labor supply as income insurance has been neglected in the literature on income risk and savings until recently.[2] Pistaferri and Saporta-Eksten (2012) and Ortigueira and Siassi (2013) are two contributions highlighting the potential importance of intra-household risk-sharing through spousal labor supply.

This paper takes a first step towards a better understanding of the nature of income

[1] Unemployment insurance does not cover 100 percent of the income and there is usually a ceiling. Evidence on credit constrained households is provided by, for example, Jappelli, Pischke and Souleles (1998), and by Crossley and Low (2011) for job losers in particular.

[2] The initial literature on the added worker effect focused on changes in labor supply along the extensive margin, see e.g. Lundberg (1985). The focus of this paper is on changes in labor supply along the intensive margin. However, we choose to label this effect the added worker-effect, as more recent literature has also used this terminology regarding the intensive margin, see e.g Pistaferri and Saporta-Eksten (2012).

risk that different households face by studying the relation between co-movement of income shocks and precautionary asset holdings. In particular, we want to assess whether households perceive spousal labor supply as a source of income insurance, and how this insurance mechanism vary with the co-movement of wage shocks. The idea is that if there is income pooling within the household and labor supply is inelastic, all that matters for the precautionary savings is the variance of total income.[3] However, if households perceive spousal labor supply as a source of insurance, it is evident that this mechanism should work better the lower the co-movement in shocks to the spousal wages. Thus, if the household is using labor supply to smooth shocks we should see that the correlation of shocks to the spouses' wages matters for precautionary wealth accumulation independent of the variance. This prediction can be tested empirically and in order to do so we use administrative data from Norway, which provides detailed information on household income, financial wealth and a rich set of control variables while minimizing the risk of measurement error.

To assess the importance of household insurance through spousal labor supply, we proceed in two steps. We first estimate the variance and the correlation of shocks to the income of individuals within a household by removing predictable changes in income and computing the variance and correlation of the residuals. More specifically, we use a rolling windows approach to estimate the parameters, which allows for the variance and the correlation to vary over time for a given household. The main problem with the obtained variables is that they are based on income and not on wages. Wage is the theoretically correct variable to study the importance of labor supply for insurance as income already incorporates labor supply movements. However, to the extent that the individuals are somewhat constrained in changing labor supply and the distribution of income shocks is similar to that of wage shocks, the correlation of income can be used to get an idea of whether labor supply is important for insurance or not. The estimates of income variance and correlation show that households in which both spouses have the same education level or are working in the same industry have a higher correlation of income shocks than couples who have different education levels and work in different industries. This is also what we would expect if assortative matching on education and industry could matter for the co-movement of shocks.

Second, we test whether households with stronger co-movement of income shocks, for a given variance, holds larger precautionary buffers. In particular, we regress different specifications of financial wealth on the variance and correlation of income shocks as well as on other relevant household characteristics. In the analysis we attempt to deal with measurement

[3]Note that the analysis only hinges on the assumption of a minimum of income pooling. That is, that the households share a strictly positive fraction of their labor income.

error in the variance and correlation estimates by using the variance in different industry combinations of the individuals in a household over time as instruments. In addition, we also attempt to control for time-invariant, unobserved differences between the households that might bias the coefficient estimates, using the fixed effects estimator. This is possible since we have obtained several estimates of the variance and the correlation for each household. The empirical results strongly indicate that intra-household insurance through labor supply is important. Our preferred estimator - the instrumental variables-fixed effects estimator - indicates that the correlation of income shocks has an important positive effect on the precautionary savings of the household, independent of the variance of total income. More specifically, we find that an increase from the lowest to the highest correlation documented in our sample would predict a median increase in financial assets of 97,144 NOK, or equivalently 16 percent of average yearly disposable income.

In our study the precautionary motive for saving is an important point of departure for the analysis. By now there are several empirical studies testing the strength of the precautionary motive by relating measures of income uncertainty to savings and asset holdings. Of these, several studies find that there is a positive relation between income risk and wealth (e.g. Guiso, Jappelli and Terlizzese, 1992; Carroll and Samwick, 1997; Kazarosian, 1997; Lusardi, 1998; Carroll, Dynan, and Krane, 2003; Alan, 2006; and Hurst, Kennickell, Lusardi and Torralba, 2006), although the importance of the precautionary motive varies between studies.[4] Our analysis shows evidence in favor of the precautionary savings hypothesis as our measure of income variance has a strong positive effect on asset holdings.

There are few studies investigating the insurance provided through individual labor supply, and even fewer studies considering the insurance provided by family members. Kotlikoff and Spivak (1981) is one of the first studies to look at insurance through family members, but they do not consider labor income. Attanasio, Low and Sanchez-Marcos (2005) model a two-person household to investigate the response of female labor force participation to idiosyncratic risks within the family while male participation is exogenous. In an attempt to study the welfare implications of a changing U.S. wage structure, Heathcote, Storesletten and Violante (2010) use a model in which both female and male labor supply is endogenous. Ortigueira and Siassi (2013) attempt to quantify the importance of intra-household risk-sharing, calibrating a model with endogenous labor supply of both spouses. Their findings indicate that this risk-sharing is important, especially for households with little wealth. Blundell, Pistaferri, and Saporta-Eksten (2012) find that, when they add family labor supply

[4]See Browning and Lusardi (1996) for a review of the earlier literature on savings choices and Meghir and Pistaferri (2011) for a summary of the more recent literature.

to assets and taxes, there is little evidence of any other insurance.

Empirical tests of the importance of spousal labor supply for income insurance have focused on the extensive margin. In particular these studies look at whether married women tend to enter the labor market in response to unemployment of their husband and find small, positive effects (e.g. Lundberg, 1985 and Cullen and Gruber, 2000; Juhn and Potter, 2007). Del Boca, Locatelli, and Pasqua (2003) do not find that the likelihood of wife employment increases with unemployment of the husband on average. However, in households where wife employment is likely to be "more accepted" they do find a positive effect. Stephens (2002) investigates the response of wives' work effort, including adjustments on the intensive margin. They find small increases in wives' labor supply prior to displacement and larger responses post displacement. Indirect evidence of the importance of spousal labor supply is provided by Browning and Crossley (2001) who find that the expenditure response to an increase in the replacement rate of unemployment insurance is higher for individuals whose spouse was not employed, indicating that spousal income is important for consumption smoothing during an unemployment spell. Similarly, findings by Lusardi (1998) indicate that the precautionary savings motive seems stronger for households with only one earner. On their part, Juhn and Potter (2007) argue that the value of marriage as insurance is likely to have decreased over time due to increased female labor force participation and a more positive co-movement of spouses' employment.

This paper contributes to the literature on insurance through spousal labor supply by considering whether households provide insurance along the intensive margin. The answer to this question has important policy implications. If spousal labor supply is an important insurance channel for households, but the ability of using it differs among households, it means that some households are to a larger extent exposed non-insurable labor market risk than others. If this risk differs between education and income groups, for example, the policy implications could be different depending on what part of the economy that is hit by a income shock. Furthermore, that households face different insurance possibilities due to different degrees of income correlation could be important to take into account when evaluating the effect of progressive income taxation and unemployment insurance. Moreover, if labor supply does provide insurance against shocks to labor income, we face the important question of to what extent unemployment insurance may crowd out labor supply, as emphasized by Cullen and Gruber (2000), Engen and Gruber (2001), and Ortiguera and Siassi (2013).

Finally, the findings of this paper points the importance of intra-household risk-sharing through labor supply for Norway, a country where the female labor force participation is high and the ability to adjust labor supply may be hampered by labor regulations. Thus we

should expect that the importance of labor supply as insurance should be even higher in a country such as the US.

The outline of the paper is as follows. The following section presents the institutional setting and the data, Section 3.3 contains the estimation of income variance and correlation. Section 3.4 offers an empirical test of the importance of spousal labor supply for insurance, and Section 3.5 concludes.

3.2 Institutional setting and data

In this section we will provide an overview of institutional details of relevance for our study, which will contribute to a better understanding of how the results for Norway may compare to other countries. More specifically we will look closer at the generosity of unemployment insurance and measures of labor market flexibility.

Unemployment insurance. Even if shocks to employment are not the focus of this paper, unemployment constitutes an important part of the income risk that the individual faces. If the state-provided unemployment insurance is a feasible alternative to household income, the importance of all insurance channels, including spousal labor supply, decreases.

Norway is among the OECD countries with the most generous welfare system and participation in welfare programs is compulsory for all residents. The most important services for adults include unemployment insurance (UI), sick money and disability pensions. The UI is an individual insurance in the sense that the benefit payment does not depend on the level of income of the spouse. There is a minimum income requirement in order to be eligible for the insurance, but this level is fairly low by Norwegian standards and in practice all full-time employees will meet this requirement. For 2007, the amount was 100,000 Norske kroner (NOK), or around 16,000 US dollar. Workers who become unemployed are entitled to around 62 percent of their earnings in the calendar year before job loss up to a ceiling.[5] The ceiling is always six times "grunnbeløpet" (the amount used as a basis to calculate social insurance payments).[6] Among the individuals in our sample in year 2010, around two thirds of the individuals earn more than the ceiling. The individual can receive UI for a maximum of 2 years.[7] When the UI expires the individual is entitled to means-tested social assistance

[5] As pointed out by Browning and Crossley (2001), a ceiling on the insurable income of an individual implies that high-income individuals will experience a larger shock to income in case of unemployment than low-income individuals as their UI replacement rate is effectively lower.

[6] "Grunnbeløpet" changes every year to account for changes in inflation.

[7] Unemployed workers above the age of 64 can receive the payments until they retire at 67.

to cover basic needs, without maximum duration.[8]

To provide a picture of the generosity of the Norwegian benefits in comparison to other countries we use calculations of benefit generosity made by the OECD.[9] Their measure of gross replacement rates compares unemployment benefits received when not working to wages earned when employed and is provided for the years 2001 to 2011.[10] According to this measure, Norway had the highest benefit ratio of the 29 countries included in 2001 at 56 percent, followed by Denmark at 54 percent. Between 2001 and 2011, the ratio decreased and the average ratio for Norway over the whole period is 47, which is the second highest average after Denmark's 50 percent. These numbers can be compared to 11 percent for the UK and 13 percent for the US. The lowest period average is that of the Czech Republic at 6 percent.

OECD also provides a similar measure[11] that is calculated for uneven years between 1975 and 2005, but may be less comparable across countries. This measure indicates that the average Norwegian replacement rate was relatively low between 1961 and 1975 (5 percent), compared to most other countries for which there is data.[12] However, it increased drastically between 1975 and 1977 from 8 to 21 percent. In 1981 Norway placed itself among the OECD countries with a relatively generous insurance. This means that the importance of individual and family insurance may have decreased more over time for a Norwegian than for individuals in other countries, as for example Denmark.

Labor market flexibility. The degree of labor market flexibility may be important for the impact of income shocks. A less rigid labor market in terms of employment protection legislation or a loser legislation when it comes to temporary work, may in itself cause greater insecurity regarding future income for those who work. On the other hand, it may make it easier to find a new job if exposed to a negative employment shock, as well as to facilitate for the partner to find a job in case of a negative shock to the partner's income.

[8]The worker could also receive sick money or disability pensions, given that a medical doctor certifies that he or she cannot work, or is permanently disabled. The maximum duration of sick money is one year and is paid at a rate of 100 percent up to the same ceiling as for UI. The replacement rate of the disability insurance is around 67 percent up the same ceiling as for UI.

[9]See Directorate for Employment, Labour and Social Affairs. Benefits and Wages: Statistics.

[10]More specifically, the benefits includes unemployment insurance and unemployment assistance benefits and the individual is the unit of analysis. The measure represents the average of the gross unemployment benefit replacement rates for two earnings levels (average earnings and two thirds of average earnings, for three different durations of an unemployment spell, and for three family and income situations (single, married with dependent spouse, and married with spouse in work). For more details, see Martin (1996).

[11]It sometimes includes also social assistance.

[12]This can be compared to for example Denmark (27), Belgium (40) and the UK (25) during the same period.

In order to see how Norway compares to other countries on these factors, we use data and studies from the OECD. OECD has a much cited measure of the strictness of the employment protection legislation (EPL) of countries. The measure is a summary index taking into account procedures and costs involved in dismissing workers. For the measure of protection of regular workers against individual and collective dismissals in 2013, Norway is at the OECD average of 2.9 out of 5. New Zealand, the US, Canada and the UK occupy the four lowest places in the ranking, while France, Netherlands, Belgium and Germany are at the top.[13]

When it comes to the use of permanent contracts, the share of dependent workers on permanent contracts rather than on temporary in Norway is on average 92.5 between 1996 and 2012. The OECD average during the same time period is 90.6 percent and for the UK 95.1.[14] In regards to the regulation surrounding temporary contracts, Norway ranks the 4th highest of the OECD countries in terms of strictness of the regulation, while UK, US and Canada hold the three bottom positions.

The extent to which it is possible to work part time is another factor likely to contribute to the flexibility of changing work hours. According to the OECD Full-time/Part-time Employment Database, the share of part-time of total employment in Norway is quite high compared to other OECD countries.[15] In 1983 it was 29.6 percent, compared to 24.8 for Sweden, and 21.2 for the Netherlands. UK had 19.0 and the US 18.4. Italy had the lowest share in this year at 4.6. The share in part-time in Norway has not followed the increasing trend seen in most other OECD countries and in 2010 it was 26.7, which is still among the highest. The only countries with higher values are Australia (29.8), Netherlands (35.2) and Mexico (27.1).[16]

To summarize, Norway seems to be a country with fairly rigid labor markets and generous unemployment insurance as compared to many other countries.

[13]OECD Employment Outlook (2013).

[14]Source OECD Labor Force Statistics. The period is from 1996 because before that data for Norway was missing. For the US, there is a lot of years missing in general.

[15]Note that the definitions of part-time varies a lot among countries. This is not of importance here however, as we want to provide a view of the individuals' possibility of changing their work hours within the national context.

[16]Even though women are contributing to most of the part-time work, also the share of males working part-time in Norway has been relatively high throughout the period, as compared to other OECD countries.

3.2.1 Data description

The data that we use is administrative data from the Norwegian tax registers between the years 1985 and 2010.[17] The dataset covers every individual in Norway and allows us to track the income of the individuals for every year since 1985.[18] Moreover, from 1993 we are able to merge this information with the financial wealth of the households through a unique personal identifier available in all registers in Norway. A couple (or a household in this context) is identified as two individuals who are married, or as two individuals who live together with common children. We can observe the latter since 1991. Unfortunately it is not possible to identify unmarried but cohabiting couples without children.

In order to estimate the variance and the correlation of job related income risk in a household we use the labor income of each individual. Labor income includes wage income and work-related transfers, such as unemployment and sickness benefits, as well as maternity leave payments. The disposable income of the individuals, i.e. the income net of taxes, is calculated for each household using the relevant tax rates for each year. In our analysis of precautionary wealth holdings the theoretically correct measure is correlation and variance of wage shocks rather than of income. Unfortunately the dataset does not contain information on labor supply for all years such that we can calculate the wages. However, even if the labor supply reacts to shocks, it may not be to the extent that the ordering of households in the income correlation and wage correlation distribution changes.[19]

For the main analysis we are interested in the relation between the variance and correlation of income risk and precautionary wealth holdings of the household (see Section 3.4). We use financial wealth as a measure of wealth.[20] One can argue that financial wealth is the most relevant wealth for the current analysis. As it is more liquid than for example housing, the individual is more likely to draw on the financial assets in times of lower income shocks than to sell the house. Chetty and Szeidl (2007), for example, argue that real estate will rarely be liquidated during unemployment spells because of high transaction costs. Basten,

[17] For more information on the Norwegian administrative data in general, see Røed and Raaum (2003) and for more information on the wealth data in particular, see Fagereng et. al. (2013).

[18] Income information is available from an earlier point in time, but financial information is only available from 1993. Why we chose 1985 will become clear with the presentation of the estimation approach.

[19] In Fagereng et al. (2014) we show that for the part of the dataset for which we have information on contractual hours there are some differences in the estimated wage and income correlations, but they are not very big. If anything, the main difference seems to be at younger ages.

[20] In the context of precautionary savings, as pointed out by Carroll and Samwick (1997) it may be preferable to use wealth accumulation rather than savings. The response to income uncertainty should generally be to hold more wealth, not necessarily to depress consumption forever. In the case of the buffer stock model, the household will not save more once the optimal stock of wealth is obtained. However, households with higher income uncertainty should hold a higher stock of wealth.

Fagereng, and Telle (2012) further argue that this is even more likely to be true in Norway due to special transaction taxes.

We define the household's financial wealth as the sum of the financial assets of the individuals in the household. The financial assets, in turn, consist of direct stock holdings, stock holdings through mutual funds, bank deposits and bond holdings. Some households have savings in pension plans, but as these plans are only payable at the age of retirement, the major part of the households in our sample will not be able to draw on these assets in times of lower income (see sample restrictions below).

The dataset further covers a rich set of characteristics for each individual such as age, education, household size, country of origin, county, and industry of the individual. We observe these variables from 1993. The industry classification that we use throughout is the 2-digit NACE code (revision 2) and for education we will divide the individuals into fairly disaggregated categories. We have the following education groups: less than secondary education; secondary education; some high school, high school degree; university lower/vocational degree or college education; university; and research. The income and asset data is deflated using "grunnbeløp i folketrygden", which is a measure accounting for the inflation used to calculate for example social benefits.

Our initial sample consists of the population of households (as defined above) in Norway. As we have data over financial assets from 1993, we are only going to use data from that year in the analysis. However, as described below our rolling window estimation will use data from 8 years back in time for each estimation. Therefore we drop all observations before year 1985. With this restriction we have 972,465 households in our sample and 14,901,340 observations, where one observation is household h in year t. Further, we drop individuals younger than 25 and older than 60 (1,393,445 observations) and we remove observations for which information is missing on county, origin, or education (198,566 observations). As the focus of the paper is on the correlation of spousal income shocks, we restrict the sample to households where both individuals have positive labor income (less 1,709,392 observations). Note that labor income also includes unemployment insurance, maternity leave benefits, etc. We further require each household to have at least 15 years of positive income observations, which need not be consecutive (less 3,334,097 observations).[21] The couples are allowed to break up and form new households which will be included if they conform to the other sample restrictions. The final sample consists of 405,888 households with an average number of periods per household of 21.0 years and a standard deviation of 3.6 years. The total number of observations is 8,265,854 household-years. The financial wealth as well as the

[21] Using only 10 years does not change the conclusions of the paper, but it adds more noise.

income will be expressed in Norwegian kroner (NOK) throughout the document. At this time one USD equals around 6.6 NOK.

The descriptive statistics for the sample used for the variance and correlation estimations are shown in Table 3.1.[22] The yearly disposable income of the females in the sample is on average lower than that of the males at 219,000 NOK compared to 351,000 NOK. The total household disposable income in the sample is on average 570,000 NOK. The average age of females is around 42.0, the average age of males is around 44.2, and the average household size is 3.8. The average number of periods (years) per households that we have for the variance and correlation estimations is 21.0 with a standard deviation of 5.4.

Table 3.1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.
Income female	219364.8	103865.3	1.1	15540156
Income male	350660.7	205418.3	1.1	145096704
Income household	570025.5	243276.5	2.4	145209904
Age female	42.0	8.2	25	60
Age male	44.2	8.2	25	60
Household size	3.8	1.0	2	6
No of years per household	21.0	5.4	15	36

3.3 Co-movements in labor income risk

3.3.1 Estimation of income variance and correlation

In the following we will estimate the income shock correlation and variance for each household.

As mentioned above, we have at least 15 non-zero observations of labor income for each individual in the household. We use the series of individual disposable incomes in a household to estimate the total variance and the correlation of household income shocks. In a first step we attempt to purge the labor income of all factors that may be predictive (to the econometrician) of the income of the individual. In order to do this we regress the natural logarithm of individual income on age and age squared, dummies for the education level, origin, county, and year, as well as the interaction of education with age and age squared and education and time.

[22] Note that the descriptive statistics of the sample used in the analysis of precautionary asset holdings will be presented in Section 3.4.2.

The regression model:

$$\ln y_{hit} = \beta_0 + \mathbf{X}_{hit}\beta_i + county_{ht} + \tau_t + \epsilon_{hit} \tag{3.1}$$

$\mathbf{X}_{hit}$ includes the observable characteristics of individual i in household h at time t, where i represents either the female or the male. τ_t is a year dummy, and ϵ_{hit} is the error term. The model is estimated by OLS and from these regressions we obtain the residuals ($\hat{\epsilon}_{hft}$ and $\hat{\epsilon}_{hmt}$).

For the estimation of the variance of total household income we use the residuals from a regression of the logarithm of total household disposable labor income (the sum of male and female disposable income) on the individual characteristics of *both* the male and the female included in the individual regressions as well as the dummies for county and time.[23]

$$\ln(y_{hft} + y_{hmt}) = \beta_0 + \mathbf{X}_{hft}\beta_f + \mathbf{X}_{hmt}\beta_m + county_{ht} + \tau_t + \epsilon_{ht} \tag{3.2}$$

It is likely that the variance and correlation vary over, for example, cohort and age.[24] We would like to capture this in our estimates of the variance and the correlation, using the household as a basis for estimation. Therefore, rather than producing one point estimate of the variance and the correlation, we estimate the variance and the correlation using a rolling window estimation in order to obtain several estimates for each household. More specifically, for each household we estimate the variance and correlation in year, t, using the income information between $(t-8:t+1)$. In order for the household to have a non-missing variance and correlation in any given year, t, the household needs a minimum of 5 years where both spouses have a positive income within the period $(t-8:t+1)$. Since we observe each household in the dataset for at least 15 years, one household should have at least 6 estimates for the variance and the correlation if none of the estimates are missing.[25]

In the sample we use for the regression analysis in the next section we drop observa-

[23] We are not controlling for household size here as we only have data on this variable since 1993. However, we are controlling for household size in the asset regressions in the next section and below we are trying to get a notion of how labor supply may contribute to the correlation patterns.

[24] Findings by Blundell, Graber and Mogstad (2013) indicate that the variance of male income vary over age, time and education. For the co-movement in income, Shore (2013) shows that it tends to increase over the time in marriage. In Fagereng et al. (2014) we find that the correlation may vary both over cohorts, age and education.

[25] We do not attempt to distinguish permanent from temporary shocks to income as it requires more information. Instead we are focusing on estimating time varying measures of the variance and the correlation for each household.

tions before 1993 as this is the year from which we have information on financial wealth of the households. With this restriction we end up with 373,209 households and 4,923,221 household-year observations. Table 3.2 displays summary statistics of the estimation results for this sample. From the bottom row we note that the average number of periods per household is around 14 years. Female income variance is higher than male income variance at 0.2 compared to 0.075. Also the standard deviation of female income variance is higher than for the male income at 0.529 compared to 0.319, although the males have a higher maximum value. The correlation is very close to zero on average with a standard deviation of 0.429.[26] The range of the estimated correlation is between -0.898 and 0.9. The average variance of total income, at 0.024, is lower than the average male variance which means that the variance of the male and the female possibly offsets each other to some extent together with the co-movements.

Table 3.2: Summary statistics variance and correlation of income shocks

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Variance female income	0.2	0.529	0	22.263
Variance male income	0.075	0.319	0	32.55
Correlation male and female income	0.001	0.429	-0.898	0.9
Variance household income	0.024	0.07	0	6.47
No of periods per household	14.063	2.891	5	17
N		4,923,221		

3.3.2 Correlation and variance over education and industry

If assortative matching on education or industry matters for the co-movement of shocks to the spousal incomes we would expect that spouses who have the same education or who work in the same industry would on average have a higher correlation, and perhaps variance, of these shocks than spouses who do not. To study whether this is the case in our sample we compare averages of the income variance and correlation of couples who have the same education level to couples who have different education levels, as well as averages of income variance and correlation of couples who work in the same industry to couples who work in different industries. The education and industry groups are represented by the male education level and industry. As mentioned before, we only observe industry since 1993 and thus, part of our total sample is not included in the comparison of industries.

[26]The findings of Shore (2013) also indicate that the average co-movement of income shocks within a household is around 0.

We start with education. The first two columns of Table 3.3 show the mean correlation over male education groups and the two latter columns display the mean variance over the same groups. Furthermore, column 1 and 3 show averages by education level for couples who have *different* education levels and column 2 and 4 show the average estimates for the households where the spouses have the *same* education level. Note that subscript h denotes household h.

First of all we may note that the differences in both variance and correlation between the male education levels are not that large in magnitude, but that the correlation on average is slightly higher for couples with the same education than for couples with different education. This difference is significant at the 0.1 percent level.[27] Furthermore, the correlation displays a u-shape over the education levels; it is highest for households where the male has a low and a high level of education and lowest when the male has a high school degree. If we compare couples with a different education level (column 1) to couples with the same education level (column 2) we notice that the average correlation is higher for couples with the same education level in all cases except for high school. However, in the case of male high school we cannot reject the hypothesis that the average correlation is the same for couples with same and different degrees. All other differences are significant at the 0.1 percent level, except for the couples with less than secondary education where the difference is not significant.

Regarding the total income variance of the household, it more or less follows the u-shaped pattern of the correlation. We may also note that the average variance for couples with same education is slightly lower than the average variance for couples with a different degree, although we cannot reject the hypothesis that they are the same. Moreover, the total variance is higher in the cases where the couple has the same education for all male education levels apart from some high school, high school and university/vocational degree. These differences are all significant. However, we cannot reject the hypothesis that the average variance is the same for the case when male has a research degree. This indicates that even if the couples with the same education tend to have a higher correlation, the average variance may be lower because of offsetting variance of the individual labor incomes.

In the next step we will do the same type of analysis over industries, which we expect to be a stronger determinant of the couple's correlation and variance of income than the education.

Table 3.4 displays the averages over industry groups. Column 1 and 2 display the correlation and column 3 and 4 show the variance. In column 1 and 3 we find averages by

[27] All the test results reported in this section are not displayed in the paper, but can be obtained by the authors upon request.

Table 3.3: Averages of variance and correlation of income shocks over male education levels

	$Mean(corr(\hat{\epsilon}_{ft}, \hat{\epsilon}_{mt}))$		$Mean(var(\hat{\epsilon}_{ht}))$	
	$educ_f \neq educ_m = j$	$educ_f = educ_m = j$	$educ_f \neq educ_m = j$	$educ_f = educ_m = j$
< Secondary	0.0303	0.0470	0.0412	0.0704
Secondary	-0.0134	0.0008	0.0248	0.0276
Some High school	-0.0044	0.0166	0.0241	0.0220
High school	-0.0052	-0.0060	0.0217	0.0209
Uni lower/vocational degree	-0.0039	0.0011	0.0239	0.0223
University	0.0101	0.0739	0.0247	0.0357
Research	0.0535	0.1180	0.0294	0.0289
Total	-0.0032	0.0073	0.0236	0.0235

Note: Column (1) and (3) refer to households which have different education level, whereas column (2) and (4) refers to households with the same education level.

industries for couples who work in different industries and in column 2 and 4 we find the average estimates for couples who work in the same industry. Although the variance estimates shown also depends on the industry of the woman, we can see from column 1 that in case the male works in some of the industries considered risky, such as agriculture and real estate activities, the total income variance is relatively high. If the male instead works in industries associated to the public sector, such as public administration and defense, as well as education, the total variance tends to be lower.[28]

We might expect that industry should be a stronger determinant of correlation and variance differences than education, but we may not that the variance and correlation differences over male industry are not that large. Comparing the average variance of couples who work in different industries to couples who work in the same industry (column 3 and 4 Table 3.4) we note from the last row of the table that, on average, the couples working in the same industry have a slightly higher total variance and the difference is significant at the 5 percent level. Furthermore, for 12 out of the 18 industries for which a comparison is possible, couples who work in the same industry have a higher variance on average, and the difference is significant at least at the 5 percent level. The exception is electricity and gas and transportation and storage, where we cannot reject that the average variance is the same between the groups. In the cases where the couples working in the same industry have a lower variance (mining and quarrying, information and communication, public administration and defense, education and human health and social work), all differences are significant at the

[28] Although not shown for this exact sample, we know from Fagereng et al. (2014) that the probability of the female working in the same industry, conditional on the male working there tend to be quite high for these industries. Therefore, the averages of total income displayed in the table can be expected to be low due to low variances in individual incomes.

0.1 percent level. It is interesting to note that half of these are industries associated to the public sector where we would expect a low variance.

When it comes to the correlation of couples working in the same industry, we see from column 2 that the highest correlation can be found in accommodation and food service activities (0.07), followed by agriculture, forestry and fishing (0.06). The lowest correlation for couples working in the same industry can be found in mining and quarrying (-0.04) and in electricity and gas (-0.02). Comparing column 1 to column 2 we see that, as expected, the average income correlation of couples working in the same industry is higher than the average income correlation of couples working in different industries for all industry groups. All differences are significant at the 0.1 percent level. The only industry where a comparison is not possible is for 'water, sewerage and waste management' where we, due to few observations, did not obtain an estimate for couples working in the same industry.

Table 3.4: Averages of variance and correlation of income shocks over male industry

	$Mean(corr(\hat{\epsilon}_{ft}, \hat{\epsilon}_{mt}))$		$Mean(var(\hat{\epsilon}_{ht}))$	
	$ind_f \neq ind_m = j$	$ind_f = ind_m = j$	$ind_f \neq ind_m = j$	$ind_f = ind_m = j$
Agriculture, forestry and fishing	0.0323	0.0599	0.0390	0.0433
Mining and quarrying	-0.0447	-0.0414	0.0201	0.0200
Manufacturing	-0.0192	-0.0044	0.0195	0.0201
Electricity and gas	-0.0289	-0.0211	0.0128	0.0136
Water supply, sewerage and waste management	-0.0069	-	0.0197	-
Construction	-0.0069	0.0020	0.0240	0.0250
Wholesale and retail trade	-0.0174	0.0059	0.0228	0.0253
Transportation and storage	-0.0124	-0.0062	0.0243	0.0250
Accomodation and food service activities	0.0119	0.0744	0.0347	0.0453
Information and communication	-0.0169	-0.0134	0.0173	0.0169
Financial and insurance activities	-0.0189	-0.0141	0.0160	0.0160
Real estate activities	-0.0076	0.0027	0.0330	0.0355
Professional, scientific and technical activities	-0.0050	0.0032	0.0237	0.0239
Administrative and support service activities	-0.0150	0.0007	0.0309	0.0333
Public administration and defense	0.0062	0.0229	0.0186	0.0182
Education	0.0088	0.0375	0.0223	0.0199
Human health and social work	0.0067	0.0350	0.0293	0.0286
Arts, entertainment and recreation	-0.0040	0.0121	0.0265	0.0275
Other service activities	-0.0147	0.0127	0.0261	0.0298
Total	-0.0109	0.0073	0.0227	0.0235

Note: Column (1) and (3) refer to households which have different industry classifications, whereas column (2) and (4) refers to households with the same industry classification.

From this descriptive analysis we conclude that working in the same industry and, to some extent, having the same education could be somewhat important for the income risk that a household faces. Thus, the industry combination of a couple is likely to be one determinant of the household's ability to use spousal labor supply as an insurance against

income shocks.

In the next section we will use our measure of the household's ability to use labor supply to insure income to test whether households consider this insurance mechanism important.

3.4 Income risk and the added worker effect

3.4.1 Prediction

In the following section we provide intuition for how the ability of an individual to insure the income of her or his spouse depends on the correlation of shocks to the individuals' incomes. Based on this reasoning we will have a prediction that we will test empirically in the following subsection.

We focus on insurance through spousal labor supply adjustments, which we refer to as the *added worker effect*. To illustrate how co-movements of spousal income risk matters for this type of insurance, imagine a household with a male and a female who both work, act as a unit and pool income within the household.[29] They both earn an hourly wage in the labor market and the wages of both individuals are subject to shocks. The shocks may be more or less correlated among the individuals depending on, for example, whether they work in the same industry or not.

In order to study the importance of co-movements of wage shocks for the added-worker-effect we will look closer at two cases: **a**) The labor supply is inelastic, i.e. the spouses cannot use spousal labor supply as insurance; **b**) Labor supply is elastic and the household can insure income through labor supply.

a. Inelastic labor supply. In this case changes in the co-movement of shocks simply affect the variability of household income. If, at the extreme, the spouses have identical income processes and a correlation of income shocks of -1, income pooling will eliminate all income risk. However, if the correlation is 1, there is no insurance through income pooling. Thus we expect that with inelastic labor supply, the co-movement of shocks to income only affects spousal insurance through the effect on total income variability.

b. Elastic labor supply. In this case the co-movement of wage shocks does not only affect total income variability, but also the effectiveness of the insurance through spousal

[29] If the individuals in the household act as two separate units with no income pooling, the correlation of labor income shocks is irrelevant for savings. Note that the analysis only hinges on the spouses sharing a strictly positive fraction of their income.

labor supply. To see this, imagine that the wage of the husband has been hit by a negative shock. If the labor market situation of the wife is also negatively affected it is likely to be more difficult for her to cover part of the income loss using her labor supply, than if she was not affected by the shock. Thus, we expect that with elastic labor supply, the co-movement of shocks to spousal wages does not only affect the spousal insurance through the effect on total income variability, but also the ability to insure income through spousal labor supply.

As such, given all other insurance channels, case **a** and case **b** have different implications for the precautionary savings of a household. More specifically, in case **a**, we would expect a household with a higher correlation to save more than a household with a low correlation of shocks, all else equal. However, given the variance of total income, the correlation of wage shocks should not matter for precautionary savings.

The situation is different when labor supply is elastic, as in case **b**. First of all, we would expect precautionary savings to be lower in this case than in case **a**, but most importantly, we would expect savings to vary with the correlation of wage shocks when we control for the variance of the wage shocks. Households with a higher correlation of shocks have a lower ability to insure income through labor supply and are expected to save more than households with a low correlation of shocks, all else equal.

Therefore, by testing this prediction empirically, we test whether spousal labor supply is considered an important insurance mechanism by the household.[30]

3.4.2 Empirical analysis

In this subsection we test the prediction stated above: If households use spousal labor supply to insure labor income, the correlation of income risk should affect precautionary savings independently of the variance in income shocks.

In order to test the prediction, we estimate regression models with accumulated financial wealth as the dependent variable and with variance and correlation of income shocks as independent variables, controlling for other household characteristics that are likely to be important for the asset holdings. If the coefficient on income correlation is positive and significantly different from zero it is evidence in favor of spousal labor supply being used to insure income.

In our analysis we use four different measures of financial wealth: the logarithm of total financial assets; total financial assets normalized by income; the logarithm of safe financial

[30]However, if the correlation is found not to have any effect independent of the total variance, we do not know whether this is because labor supply is inelastic of because spouses are not pooling their income.

assets; and safe financial assets normalized by income. The safe asset category excludes direct and indirect stock holdings and thus it consists of bank deposits and bond holdings. To the extent that safe assets are more liquid than risky assets we may see a difference in the effect of income risk only including these assets.

As mentioned before, we observe financial assets of households from 1993. Given the procedure to estimate the variance and the correlation of income risk of the households using a ten period rolling window from $t-8$ to $t+1$, we will with our sample from 1985 have the first variance and correlation estimates for year 1993. Since we will not use observations before 1993 in the regressions we drop these observations and are left with 373,209 households and 4,923,221 observations in total.

The approach we use to estimate income uncertainty assumes that households are mainly backward-looking when forming expectations of the future variance and correlation. In our view this is a reasonable assumption and it leaves us with more observations per household to estimate the effect of income uncertainty on financial asset holdings.[31]

The descriptive statistics of the sample we use for the regressions are presented in Table 3.5. The average (disposable) household income in the sample is around 600,000 NOK. The average financial wealth per household in the sample is around 456,000 NOK. The ratio of financial wealth to disposable income is on average 0.75, but we see that maximum value is 35, indicating that most individuals in the sample has a low financial wealth to income, but that there are exceptions. The average of safe assets is around 214,000 NOK. As the minimum value of this variable is 0 we can conclude that the entire financial wealth of at least one of the households consists of stocks. The average household observation has a higher yearly income than safe asset holdings as the average is around 0.51. The average age of males is around 46 years old and of females around 44. The average household size in our sample is 3.7 individuals. We also see from the table that in 36 percent of the households the individuals have the same education category and in 18 percent of the households the individuals have the same industry classification (2-digit NACE, rev. 2).

We estimate the following regression model:

$$\ln W_{ht} = \alpha_0 + \alpha_1 \ln Var_{ht} + \alpha_2 Corr_{ht} + \alpha_3 \mathbf{X}_{ht} + \theta_t + \eta_h + \nu_{ht} \tag{3.3}$$

where W_{ht} is one of the four measures of financial wealth, $\ln Var_{ht}$ is the logarithm of our measure of the variance of income shocks and $Corr_{ht}$ is the measure of the correlation of

[31] One potential problem with the current approach is that we assume that all changes that affect the income uncertainty are unpredictable shocks, as for example a change of industry. However, using only future periods to estimate the uncertainty we assume that they are all predicted.

Table 3.5: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.
Household income	600265.0	162468.4	224003.1	1330590
Financial wealth	456416.5	848761.1	236.8	8339515
Ratio of financial wealth to disp. income	0.754	1.421	0.001	35.23
Safe assets	214243.7	395410.3	0	8010357
Ratio safe assets to disp. income	0.506	0.901	0	30.024
Age, male	46.4	7.4	28	60
Age, female	44.3	7.5	28	60
Household size	3.7	1.1	2	6
Share in same education	0.357	0.479	0	1
Share in same industry	0.184	0.387	0	1
Variance income	0.024	0.07	0	6.47
Variance female income	0.2	0.529	0	22.263
Variance male income	0.075	0.319	0	32.55
Correlation income	0.001	0.429	-0.898	0.9
N		4923221		

income shocks. $\mathbf{X}_h$ is a vector of household characteristics including the logarithm of the total household disposable income, male age and age squared, dummies for household size, and for the industry and education of both spouses, as well as year dummies. Moreover, households may differ in some unobserved variable, which might also be correlated with the outcome variable. This variable is denoted by θ_h and we assume that θ_h is time invariant.

In a first step we estimate this model using the ordinary least squares (OLS) estimator. However, the OLS estimator is likely to be problematic for this analysis for several reasons. First of all, as we are using at most 10 observations for each estimate of a household's variance and correlation of income, the household variance and correlation are likely not to be precisely estimated. Thus, we have a problem of measurement error in the variables of interest and, assuming that the measurement error is classical, the estimated coefficients are downward biased. In an attempt to correct for this we follow the approach of, for example, Carroll and Samwick (1997) and use the instrumental variables (IV) estimator.[32] As instruments we will use the variation in couple combinations of industry over time. For example, all males that work in education with a spouse working in construction, according to the 2-digit NACE classification will be one combination, and so forth. More specifically,

[32]However, we go about it slightly different with respect to their study. In their analysis, Carroll and Samwick (1997) use industry, education, and occupation for the male as well as age interactions of these variables as instruments.

the instruments are the average income shock correlation and variance for each such industry combination at each point in time. We consider this to be an appropriate instrument as the industry combination of a couple is likely to be an important determinant of the income risk and, in addition, the influence of such a combination may change over time. The exclusion restriction is that the industry variances and correlations estimated for each spousal industry combination do not influence the financial wealth of the household other than through the household income variance and correlation. We consider this to be a plausible assumption as it is difficult to imagine how the couple combination of industry would influence the financial wealth beyond the effect on income variance and correlation once we control for the industries of the individuals. An additional benefit of using the instruments is that we might take care of time-varying factors that are specific to the household and that may affect both the uncertainty estimates and the financial buffer, like a health shock for example. Using the instruments we attempt to capture the effect on the household's precautionary asset holdings that stem from the uncertainty expectations the household has by belonging to a specific industry combination group over time.

Another potential problem with the OLS estimator is that households with a high income shock correlation or variance could be systematically different from households that have a low income correlation or variance in ways that we are not able to control for and that may be relevant for precautionary savings. That is, we have an omitted variables problem. As mentioned above, one example would be risk aversion. Households with a lower risk aversion may choose to work in more risky industries (as well as the same industry), and they may save less.[33] In order to deal with the problem of unobserved heterogeneity we assume that this unobserved household-specific factor is fixed over time and we account for it using the fixed effects (FE) estimator.

Finally, in order to try to correct both endogeneity problems at the same time, we use a fourth estimator, the IV-FE estimator, which is our preferred estimator.

First stage

Before we go to the main results we show the results from the first stages. Due to measurement error, as well as other potential problems, in our variables of interest we have two endogenous regressors in the asset regressions. In the first stages of the IV and the IV-FE estimations we predict each of these regressors using both our instruments. As stated above,

[33]Shore (2013) finds that couples seem to display positive assortative matching on income risk, in the sense that a husband with a higher income volatility is more likely to have a wife with higher income volatility. This finding points to the conclusion that individuals match on risk aversion.

the instruments we use are the average variance and correlation of income shocks over male and female industry combinations and year, denoted by $VarInd_{ct}$ and $CorrInd_{ct}$, where c denotes industry combinations (i.e. male works in education and the female in construction, etc.). Thus, the first stage equations are:

$$\ln Var_{ht} = \gamma_0 + \gamma_1 \ln VarInd_{ct} + \gamma_2 CorrInd_{ct} + \gamma_3 \mathbf{X}_{ht} + \lambda_h + \omega_{ht} \tag{3.4}$$

$$Corr_{ht} = \delta_0 + \delta_1 \ln VarInd_{ct} + \delta_2 CorrInd_{ct} + \delta_3 \mathbf{X}_{ht} + \kappa_h + \varpi_{ht} \tag{3.5}$$

The results from the first stages are displayed in Table 3.6. In column 1 and 2 we show the results from the first stage of the IV. We can see that our instrument for correlation is a strong predictor of the couple correlation and, similarly, from column 2 we note that the instrument for variance is a strong predictor for the variance in income shocks of the couples. We may note from the table that the F-values are very large, indicating that the instruments are strong. In column 3 and 4 we show the results from the first stage within regression.[34] We can see that the results are similar to those in column 1 and 2, although the effect of the instruments is somewhat lower. The F-values of the instruments are still very high. Thus, these first stage results seem to indicate that our instruments are strong and we will proceed to the main analysis of the paper

The effect of variance and correlation on precautionary wealth holdings

In the following we test whether spousal labor supply (or the added worker effect) is an important labor income insurance mechanism for households in Norway. As stated above, we will use four different outcome variables: log financial wealth, the ratio of financial wealth to disposable income, the log of safe financial assets, and the ratio of safe financial assets to disposable income. Before the regressions we winsorize the financial asset categories at the 99 percentile. This is also why the number of observations are differing slightly between the regressions of different asset categories.

The first table, Table 3.7, shows the results using log financial wealth of the household as the dependent variable. The first column displays the results of the OLS estimator from a regression of financial wealth on the logarithm of income variance and correlation. The coefficient on income variance is positive and significant, as we would expect from earlier studies. The correlation is also positive and significant, but it has a very small impact. In column 2 we see the results from the second stage of the IV regression. The impact of variance

[34]Note that the education dummies drop out in the within estimation as education is fixed over time.

Table 3.6: First Stage IV and IV-FE

	First stage IV		First stage IV-FE	
	(1)	(2)	(3)	(4)
	Corr_{ht}	$\ln Var_{ht}$	Corr_{ht}	$\ln Var_{ht}$
VarInd_{ct}	0.001	0.693***	-0.002	0.380***
	(0.00167)	(0.00518)	(0.00168)	(0.00430)
CorrInd_{ct}	0.999***	0.075**	0.613***	-0.015
	(0.00830)	(0.02579)	(0.00914)	(0.02322)
Log disposable income	yes	yes	yes	yes
Age polynomial	yes	yes	yes	yes
Education dummies	yes	yes	no	no
Industry dummies	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes
Household size	yes	yes	yes	yes
F-value (instruments)	9374	11704	2598	4578
N	4923221	4923221	4923221	4923221

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

is now stronger. The coefficient on correlation is much stronger with this specification and indicates that the added worker effect is important for Norwegian households. However, we would still be worried about potential omitted variables bias. In column 3 we attempt to deal with this problem using the fixed effects estimator. We see that both the coefficient on the variance and the correlation becomes smaller with respect to both the IV and the OLS estimator, indicating that there might be an unobserved, household-specific variable that is positively correlated with income risk and precautionary savings. When we attempt to deal with both unobservable variables and measurement error using our preferred estimator, the IV-FE estimator (column 4), we see that both the variance and the correlation of couples' income shocks have an important impact on precautionary financial wealth of the household. A one percent change in the variance leads to 0.31 percent increase in financial wealth and one unit change in the correlation increases the financial wealth by about 42 percent[35]. In regards to the correlation impact, moving from the lowest (-0.9) to the highest correlation (0.9) in our sample is predicted to give a median increase in financial assets of 97,144 NOK, or equivalently 16 percent of yearly disposable income.[36] Increasing the total variance with

[35] $42 = 100 * (exp0.351 - 1)$

[36] The correlation impact is found by predicting the financial wealth for all households, using the IV-FE estimates, assuming the all households have the highest and the lowest correlation. Then, we compute the

one standard deviation, as reported in Table 3.2, is predicted to give a median change in financial wealth of 146,491 NOK, corresponding to about 24 percent of yearly disposable income. Note however that the standard deviation of the total variance is about three times larger than the average value.

Table 3.7: Regressions with financial wealth as dependent variable

	(1) OLS	(2) IV	(3) FE	(4) IV and FE
$\ln Var_{ht}$	0.120***	0.566***	0.0575***	0.307***
	(0.000534)	(0.0101)	(0.000508)	(0.0115)
$Corr_{ht}$	0.00621***	1.613***	-0.0391***	0.351***
	(0.00165)	(0.0353)	(0.00127)	(0.0396)
$\ln Income_{ht}$	1.077***	1.766***	0.373***	0.662***
	(0.00308)	(0.0133)	(0.00297)	(0.0119)
Age 2nd polynomial, male	yes	yes	yes	yes
Household size	yes	yes	yes	yes
Education dummies (f,m)	yes	yes	no	no
Industry dummies (f,m)	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes
N	4923221	4923221	4923221	4921922

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.8 shows the results using financial wealth normalized by income as dependent variable. The results show a similar pattern to that of financial wealth. The variance is positive and significant for all specifications with a stronger impact in the IV specifications. The results in column 4 indicate that a one percent increase in income variance increases the precautionary savings, measured by the ratio of financial wealth to income, by around 0.36. Regarding the correlation, it is now negative with a small effect when estimated by OLS, but as before the impact is much stronger, and positive, with the IV estimator. In the fixed effects estimation, correlation is again estimated to have a stronger negative impact relative to the OLS estimator. With our preferred estimator, however, we see that the correlation has a strong, positive impact, independent of the variance, on precautionary wealth holdings measured by the share of financial wealth to income. A unit change in the correlation results in an increase of 0.15 in the ratio of financial wealth to income.

median change in financial wealth between the two predicted values.

Table 3.8: Regressions with ratio of financial wealth to income as dependent variable

	(1) OLS	(2) IV	(3) FE	(4) IV and FE
$\ln Var_{ht}$	0.109***	0.539***	0.0430***	0.362***
	(0.000386)	(0.00809)	(0.000346)	(0.00871)
$Corr_{ht}$	-0.0227***	1.381***	-0.0311***	0.148***
	(0.00107)	(0.0272)	(0.000863)	(0.0294)
$\ln Income_{ht}$	0.0659***	0.704***	-0.509***	-0.173***
	(0.00248)	(0.0106)	(0.00204)	(0.00901)
Age 2nd polynomial, male	yes	yes	yes	yes
Household size	yes	yes	yes	yes
Education dummies (f,m)	yes	yes	no	no
Industry dummies (f,m)	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes
N	4824755	4824755	4824755	4822979

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In Table 3.9 we use the logarithm of safe financial assets as the dependent variable and from column 1 we see that the pattern is similar also for this wealth measure. We observe small effects from the OLS regression, and a more negative impact when controlling for a time invariant unobservable variable in column 3. In the IV estimation the effects are stronger and using our preferred estimator we obtain strong, positive effects for both variance and correlation, although we see that the effect is smaller than for total financial assets. A one percent increase in the variance of income shocks result in a 0.24 percent increase in safe asset holdings and a unit change in the correlation results in an increase of around 21 percent in safe asset holdings.

Finally, in Table 3.10 we show the results from a regression using safe assets normalized by household disposable income. The same patterns as for the former regressions using the other outcome variables show up here, but with a lower magnitude of the coefficients. From column four we note that a one percent increase in the variance of income shocks leads to an increase of the ratio of safe assets to income of 0.15. The impact of the correlation is also much lower, but still positive and significant. A one unit increase in the correlation leads to a 0.07 increase in the same ratio.

From the analysis presented above we conclude that a higher variance in shocks to household income increases the precautionary asset holdings of a household. Importantly for our

Table 3.9: Regressions with safe assets as dependent variable

	(1) OLS	(2) IV	(3) FE	(4) IV and FE
$\ln Var_{ht}$	0.0781***	0.334***	0.0617***	0.249***
	(0.000574)	(0.00941)	(0.000608)	(0.0133)
$Corr_{ht}$	0.0325***	0.928***	-0.0421***	0.190***
	(0.00173)	(0.0327)	(0.00152)	(0.0463)
$\ln Income_{ht}$	0.772***	1.166***	0.326***	0.538***
	(0.00340)	(0.0124)	(0.00356)	(0.0137)
Age 2nd polynomial, male	yes	yes	yes	yes
Household size	yes	yes	yes	yes
Education dummies (f,m)	yes	yes	no	no
Industry dummies (f,m)	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes
N	4923221	4923221	4923221	4921922

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.10: Regressions with ratio of safe assets to income as dependent variable

	(1) OLS	(2) IV	(3) FE	(4) IV and FE
$\ln Var_{ht}$	0.0496***	0.186***	0.0281***	0.152***
	(0.000235)	(0.00404)	(0.000238)	(0.00549)
$Corr_{ht}$	-0.00142*	0.475***	-0.0206***	0.0686***
	(0.000673)	(0.0137)	(0.000593)	(0.0186)
$\ln Income_{ht}$	-0.110***	0.0927***	-0.353***	-0.222***
	(0.00149)	(0.00525)	(0.00141)	(0.00560)
Age 2nd polynomial, male	yes	yes	yes	yes
Household size	yes	yes	yes	yes
Education dummies (f,m)	yes	yes	no	no
Industry dummies (f,m)	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes
N	4824755	4824755	4824755	4823128

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

study, we find that the correlation of income shocks within the household has a positive impact on precautionary savings, controlling for the variance in income shocks. This indicates that spousal labor supply is an important labor income insurance channel of the households in Norway.

Additional checks

In order to make sure that the results are robust to slightly different specifications of the model we have made some additional estimations.

In the current regressions we have included the correlation of individual income risk. The correlation requires more information to estimate than the covariance and we might be worried that the measurement error problem is larger for the correlation. This should effectively be taken care of by the instruments, but in order to show that our main conclusion does not change using the covariance we display the results of the IV-FE estimator using the logarithm of total financial wealth as the outcome variable. The results are displayed in column 1 of Table 3.11 and we can see that both variance and covariance are positive and significant, confirming the main conclusion from above.

Another concern may be that the households who hold very little safe assets are different in ways that could influence the regression results. To make sure that this is not the case, we check the results of the IV-FE estimator using the logarithm of total financial wealth as the outcome variable, restricting the estimation to households with at least 3000 NOK in safe assets. The result is shown in column 2 of Table 3.11. As we can see the results are very similar to what is found above.[37]

The conclusion from this analysis is that the correlation of income shocks within a household has an independent, positive impact on precautionary asset holdings of the household, controlling for income variance. As our theoretical prediction seems to hold in the data we further conclude that spousal labor supply *is* important for the income insurance of a household.

[37]Importantly, Hurst, Kennickell, Lusardi and Torralba (2006) note that including business owners in the analysis increases the importance of the precautionary savings motive. The reason is that business owners tend to have a higher income variance and hold higher wealth. We have also run the regressions without self-employed individuals and it does not change the results.

Table 3.11: Additional checks

	(1) Safe assets > 3000	(2) Covariance
$\ln Var_{ht}$	0.322***	0.295***
	(0.0109)	(0.00966)
$Corr_{ht}$	0.339***	
	(0.0371)	
Cov_{ht}		3.415***
		(0.374)
$\ln Income_{ht}$	0.645***	0.607***
	(0.0111)	(0.0108)
Age 2nd polynomial, male	yes	yes
Household size	yes	yes
Education dummies (f,m)	no	no
Industry dummies (f,m)	yes	yes
Year dummies	yes	yes
N	4799996	4921922

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3.5 Conclusion

In this paper we investigate the importance of intra-household risk-sharing through labor supply, or the so called added worker effect. More specifically, we test the theoretical prediction that, in case households are using labor supply to smooth income shocks, the correlation of these shocks should have a positive and significant effect on savings, controlling for the variance of shocks to total household income.

In order to test this prediction we use administrative data from Norway. In a first step we estimate the variance and correlation of income shocks for each household. We find that the correlation of income shocks in the population is very close to zero on average, but that there is some important variation. In particular, couples who work in the same industry have a higher correlation of shocks than those who do not, which we would expect if the estimated correlation could tell us something about the common risk that these couples face.

Importantly, in the main analysis we find that the variation in correlation across households is not trivial from the perspective of income insurance. The results from our regression

analysis show that the correlation has an impact on precautionary asset holdings, independent of the variance of income shocks. In the case where all households move from the lowest to the highest correlation documented in our sample, the predicted median change in asset holdings is 97,144 NOK per, which corresponds to 16 percent of yearly disposable income. This suggests that spousal labor supply, or the added worker effect, is indeed important for income insurance and thus that the co-movement of income risk is important for the household's ability to insure income. This may have important implications for the design of tax and benefit systems, such as the unemployment insurance. Furthermore, it is important to emphasize that our findings relate to Norway, a country with a relatively rigid labor market and a generous welfare system. We therefore expect that our results provide a lower bound on the importance of insurance through spousal labor supply and we would expect the importance to be even greater in countries with more flexible labor markets and less generous welfare systems, as for example the US.

Bibliography

[1] Alan, S. (2006). Precautionary wealth accumulation: Evidence from Canadian microdata. *Canadian Journal of Economics, 39*, 1105-1124.

[2] Attanasio, O., Low, H. & Sanchez-Marcos, V. (2005). Female labor supply as insurance against idiosyncratic risk. *Journal of the European Association, 3*, 755-764.

[3] Basten, C., Fagereng, A. & Telle, K. (2013). Saving and portfolio allocation before and after job loss. Working paper.

[4] Blundell, R., Graber, M. & Mogstad, M. (2014). Labor income dynamics and the insurance from taxes, transfers and the family. *Forthcoming in Journal of Public Economics.*

[5] Blundell, R., Pistaferri, L. & Saporta-Eksten, I. (2012). Consumption inequality and family labor supply. NBER Working Paper, No. 18445.

[6] Browning, M. & Crossley, T.F. (2001). Unemployment insurance benefit levels and consumption changes. *Journal of Public Economics, 80*, 1-23.

[7] Browning, M. & Lusardi, A. (1996). Household saving: Micro theories and micro facts. *Journal of Economic Literature, 34*, 1797-1855.

[8] Carroll, C., Dynan, K.E. & Krane, S.S. (2003). Unemployment risk and precautionary wealth: evidence from households' balance sheets. *Review of Economics and Statistics, 85*, 586-604.

[9] Carroll, C.D. & Samwick, A.A. (1997). The nature of precautionary wealth. *Journal of Monetary Economics, 40*, 41-71.

[10] Chetty, R. & Szeidl, A. (2007). Consumption commitments and risk preferences. *The Quarterly Journal of Economics, 122*, 831-877.

[11] Crossley, T. & Low, H. (2011). Jobb loss, credit constraints and consumption growth. Mimeo, University of Cambridge.

[12] Cullen, J.B. & Gruber, J. (2000). Does unemployment insurance crowd out spousal labor supply? *Journal of Labor Economics, 18*, 546-572.

[13] Del Boca, D., Locatelli, M. & Pasqua, S. (2003). Employment decisions of married women: Evidence and explanations. *Labour, 14*, 35-52.

[14] Engen, E.M. & Gruber, J. (2001). Unemployment insurance and precautionary saving. *Journal of Monetary Economics, 47*, 545-579.

[15] Fagereng, A., Gottlieb, C. & Guiso, L. (2013). Asset market participation and portfolio choice over the life-cycle. Netspar Working Paper.

[16] Fagereng, A., Gustafsson, M. & Naess Torstensen, K. (2014). Co-movements in household income shocks. Mimeo.

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

[18] Heathcote, J., Storesletten, K. & Violante, G.L. (2010). The macroeconomic implications of rising wage inequality in the U.S. *Journal of Political Economy, 118*, 681-722.

[19] Hurst, E., Kennickell, A., Lusardi, A. & Torralba, F. (2006). Precautionary savings and the importance of business owners. CFS Working Paper 2006/16.

[20] Jappelli, T., Pischke, J-S. & Souleles, N. (1998). Testing for liquidity constraints in Euler equations with complementary data sources. *Review of Economics and Statistics, 80*, 251-262.

[21] Juhn, C. & Potter, S. (2007). Is there still an added worker effect? Staff Report, Federal Reserve Bank of New York. No. 310.

[22] Kazarosian, M. (1997). Precautionary savings - a panel study. *Review of Economics and Statistics, 79*, 241-247.

[23] Kotlikoff, L.J. & Spivak, A. (1981). The family as an incomplete annuities market. *Journal of Political Economy, 89*, 372-391.

[24] Lundberg, S. (1985). The added worker effect. *Journal of Labor Economics, 3*, 11-37.

[25] Lusardi, A. (1998). On the importance of the precautionary savings motive. *American Economic Review, 88*, 449-453.

[26] Martin, J. (1996). Measures of replacement rates for the purpose of international comparisons: A note. *OECD Economic Studies, No. 26.*

[27] Meghir, C. & Pistaferri, L. (2004). Income variance dynamics and heterogeneity. *Econometrica, 71*, 1-32.

[28] Meghir, C. & Pistaferri, L. (2011). Earnings, consumption and lifecycle choices. *Handbook of Labor Economics, 4*, 773-854.

[29] Ortigueira, S. & Siassi, N. (2013). How important is intra-household risk sharing for savings and labor supply? *Journal of Monetary Economics, 60*, 650-666.

[30] Roed, K. & Raum, O. (2003). Administrative registers - unexplored reservoirs of scientific knowledge? *The Economic Journal, 113*, F258-F281.

[31] Shore, S.H. (2013). The co-movements of couples' incomes. *Review of Economics of the Household*, May 2013.

[32] Stephens Jr., M. (2002). Worker displacement and the added worker effect. *Journal of Labor Economics, 20*, 504-537.