UNIVERSITÀ COMMERCIALE "LUIGI BOCCONI" – MILAN
Faculty of Economics
Institute for Quantitative Methods
Ph.D. Dissertation in Statistics
XIX Cycle

Modeling Fertility Hazards with Unobserved Heterogeneity

Supervisor:
Prof. Francesco C. BILLARI

Ph.D. Dissertation by:
CONCETTA RONDINELLI
Matr. n°: 935692

Anno Accademico 2005-2006

UNIVERSITÀ COMMERCIALE “LUIGI BOCCONI”
INSTITUTE FOR QUANTITATIVE METHODS

The thesis “**Modeling Fertility Hazards with Unobserved Heterogeneity**” by **Concetta Rondinelli** is recommended for acceptance by the members of the delegated committee, as stated by the enclosed reports, in partial fulfillment of the requirements for the degree of **Doctor of Philosophy**.

Dated: January 2007

Research Supervisor: **Prof. Francesco C. BILLARI**

External Examiners: **Prof. Arnstein AASSVE**

Prof. Michele PELLIZZARI

UNIVERSITÀ COMMERCIALE “LUIGI BOCCONI”

Date: **January 2007**

Author: **Concetta Rondinelli**

Title: **Modeling Fertility Hazards with Unobserved Heterogeneity**

Department: **Institute for Quantitative Methods**

Signature of Author

To my Parents.

Table of Contents

List of Tables

List of Figures

Abstract

The aim of this thesis is to study fertility using hazard modeling and controlling for unobserved heterogeneity.

In order to assess the impact of socio-economic status on fertility decisions, we focus on the period during which lowest low fertility emerged in Italy. We start by constructing a data set which makes use of two different surveys: one to infer socio-economic status, from Bank of Italy (SHIW), the other to estimate models of first, second and third births, from ISTAT (Labor Force Survey). Starting from a basic single spell discrete time hazard model without Unobserved Heterogeneity, we extend it to a more complex multiple spell duration model that controls for women specific unobserved characteristics. We show that women's socio-economic status indeed delays the onset of motherhood, but also that there are strong recuperation effects which works through first and second births. Finally, we evaluate, with the help of Monte Carlo simulations, the consequences of omission or misspecification of the unobserved heterogeneity in discrete time duration models with single spells. The findings are encouraging for practitioners who would like to estimate discrete time duration models by using sequential discrete binary models with or without Gaussian random effects which are implemented in standard econometric software packages.

JEL Classifications: J13, J18, C41, C24, C15.

Keywords: Discrete Time duration models, Spline, Unobserved Heterogeneity, Fertility, Italy, Lowest-Low fertility, Postponement and Recuperation Effect.

Acknowledgements

I would like to thank Prof. F. C. Billari, my supervisor, for his constant support and many suggestions during this research. He is simply brilliant and his ideas were crucial to the successful completion of the thesis. His guidance was fundamental for my studies on socio-economic analysis. He really is world class! From his passion for demographic methods I learnt that Statistical Methods should be used to implement new policies and give advice to policymakers. Again, he is simply brilliant!

The Institute for Social and Economic Research (ISER), University of Essex, expressed its interest in my work. Thanks to Prof. J. Ermisch and Prof. S. Jenkins for their helpful comments and suggestions. Dr. A. Aassve initialized myself to the data management: thanks to his patience and advices I improved my research. I am grateful to him because we started from basic concepts and we ended up with more advanced and innovative topics on hazards models. Dr. C. Nicoletti gave me a good perspective about methodological skills on duration models. From her I learnt how statistical subjects should be treated in a formal way. I should also mention all the participants to the Demographic Workshop in May 2006 at ISER for their comments on my first chapter, all the ISER staff for their hospitality during my research visit and, of course, all the Italian group at ISER.

Thanks to Prof. Muliere for addressing myself to Statistical Methods and to Prof. Cifarelli for his helpful guidance and advices when taking difficult decisions. I am also grateful to Dr. R. Graziani for her guidance through the continuous moments of chaos and confusion.

Of course, I am grateful to my parents for their patience and for encouraging myself to continue my Ph.D. in Statistics in difficult moments.

Thanks to Pasquale for his love and for supporting myself through the year spent in the UK, for all the travels to visit me in these three years and for all the week-ends spent together studying.

I wish to thank my brothers Nico, Enzo and Pierfrancesco, because they are special, and all the rest of my family for encouraging my choices. Finally thanks to my friends: Raffaella, Francesca, Giulia, Aurora, Catia, Silvia, Valeria, Glenda, Jacopo, Elena, Ippolita.

Milan, Italy Concetta Rondinelli

January 20, 2007

Introduction

The modeling of waiting times to particular events experienced by, for instance, individuals or firms, has become a popular and well established topic in statistics. Information on waiting times, normally referred to as durations, are readily available in numerous data sources and used in a wide array of applications.

Statistical models dealing with durations have a long history. The first examples of applications were in mortality analysis and clinical trials, where the duration of interest represents the length of time between births, diagnosis or beginning of a medical treatment, or death; and in reliability analysis, where the duration represents the length of time from the first use of an item to its mechanical failure. Those first applications have influenced the terminology used in the statistical literature. The event ending a duration is called risk or failure event; the probability that an individual has not yet experienced the risk event at a specified time (duration) is termed survival probability; and the instantaneous probability that an individual experiences the risk event at a specific time, given that she (he or it) did not experience it earlier, is termed the hazard function. There are numerous examples of duration analysis in economics and social sciences, but where the failure or risk events tend to refer to positive events such as finding a job (Lancaster 1979) or giving births (Newman and McCulloch 1984).

One of the points that has been stressed by econometricians (and that was originally neglected by biometricians) is the fact that statistical models are often incomplete. That is, even if an attempt has been made to include all the relevant factors determining the outcome of interest, there tend to be important variables that are not available in the data, and therefore not included in the model. These omitted variables are generally referred to as *Unobserved Heterogeneity.*

In this work we focus on the way in which fertility hazards are modeled. More precisely, we study the consequences of ignoring the Unobserved Heterogeneity in duration models, both from a methodological perspective and from the applied side. We also present specific analyses of the Italian case.

The thesis is organized in 3 chapters. In Chapter 1, we propose a basic hazard specification for the analysis of fertility choices using Italian data. Chapter 2 describes a non standard hazard model with Unobserved Heterogeneity in a multi spell context and discusses the empirical evidence concerning postponement and recuperation of fertility in Italy. Chapter 3 concludes with methodological issues and consequence of ignoring or misspecifying the Unobserved Heterogeneity in discrete time duration models with single spell.

Chapter 1 deals with the relationship between women's wages and childbearing decisions in Italy.

It is a joint work with Arnstein Aassve and Francesco C. Billari: part of the chapter has been published as an Institute for Social and Economic Research (ISER), University of Essex, working paper (n. 2006-06). This work was conducted during a research visit in 2005 and

2006, and the author gratefully acknowledges the hospitality and support provided by ISER. The paper was presented by me at the European Society Population Economics (ESPE, June 2006, Verona) and at the European Economic Association Meeting (EEA, August 2006, Vienna). A range of useful comments were provided during these presentations. The chapter also benefits from the comments of the participants to the Demographic Workshop in May 2006 at ISER.

The chapter assesses the impact of women's wages on childbearing decisions: so far, research has been hampered by the availability of data combining fertility and wage. This is especially a problem for third birth parity. We overcome this problem by combining two different data sets. Data from the Survey of Italian Households' Income and Wealth (Bank of Italy 2002) are used to predict women's potential wage, which is then introduced as explanatory variable in discrete time hazard regression models for first, second and third births. Data from the Labor Force Survey (ISTAT, 2003) are used in order to reconstruct women's demographic and fertility histories. Since we use wages predicted from the Bank of Italy data set as covariates in the ISTAT data, the standard errors might have an upward bias. To overcome this limit we bootstrapped the standard errors using 300 replications with observations clustered at the woman level.

We start by modeling a simple hazard specification without Unobserved Heterogeneity, with single spells and set in discrete time. The work is innovative because of 1) the linkage of different data sets and 2) because there are no available studies on the role that wage plays on the transition to first, second and third birth for the Italian context. The wage effect is negatively associated with childbearing, though the magnitude varies according to birth parity. Consistent with opportunity cost theory, wage has a strong and negative effect in the timing of first birth and women with higher wages tend to delay motherhood. We find that wage has a non-proportional effect on the hazard, and that there is a "recuperation effect" (even if not complete), the latter suggesting that women with higher wages start to have children later, but recuperate after some time. Furthermore, there is no evidence that institutional effects are responsible for the postponement of maternity. Different is the pattern for second and third birth: the wage effect is negative but it has smaller intensity when compared to the first one. Nevertheless, in line with Ermisch (1989), we find evidence of institutional effects affecting the decision of having more than one child: the risk of experiencing a second and third birth is higher for a Northern woman because she is more confident in the availability of childcare.

It is reasonable to suppose that increasing financial support to households with children may have an impact on the probability of having the third and fourth child, mainly for poorer households. To this extent, the Welfare Minister Livia Turco (law number 448 of the Year 1998) introduced two policy measures with the explicit purpose of supporting poor households with children. These two measures could provide a significant increase in income for low-income families, covering a non-negligible proportion of the cost of an additional child. Our estimates give some support for the law having an impact. But it is not clearly identifiable and quantifiable given our methodological approach, which in turn is caused by the data available. Despite this limitation our results provide interesting evidence of the key role wage plays in fertility decisions.

Chapter 2 is based on the data constructed in the first chapter. The specification is a multi-spell random effect hazard model, whereby estimation of first, second and third birth is done jointly. The model allows for unobserved heterogeneity and we test to what extent the assumption about its functional form has any impact on the overall parameter estimates.

This chapter is part of the working paper published at ISER (n. 2006-46) and it is coauthored with Arnstein Aassve and Francesco C. Billari.

The research is innovative since we model fertility hazards with a Gaussian Unobserved Heterogeneity and for then to use a Non Parametric Maximum Likelihood Estimation (NMPLE) to assess the number of relevant masspoints.

Using the two data sets of Chapter 1, one to infer socio-economic status, the other to estimate models for first, second and third birth, a bootstrap estimation of the standard errors in the hazard equation is needed because of the imputed regressors. All analysis of this chapter have been conducted with a free package aML, an useful software for studying multi-spell, multi-level models with huge data sets. aML does not have a pre-built procedure for the bootstrap method and we used a PERL routine to do it. We would like to thank Dr. Fabrizio Iozzi for helping us in writing and implementing this procedure. At each replication, we bootstrapped the 20% of the original sample. We used 100 replications and clustered the observations at the woman level.

One of the major explanations brought forward for the emergence if low fertility in Italy (and other Mediterranean countries) has been increased education and labour force participation among women. The argument relies on the fact that higher education and earnings increase women's opportunity cost, which in turn delay the onset of childbearing and therefore reduce completed fertility. As expected, our model shows that women with high wages tend to delay the first birth: this is not unexpected since education is the most powerful predictor for potential earnings. However, we also find a strong recuperation effect of high earning women. The recuperation takes place both through first and second births. As a consequence, by the age of 40 high earning women have caught up with low earning women almost completely. Interestingly, socio-economic status as measured by the predicted wage, has little effect on third birth.

To overcome the estimation problems caused by women specific unobserved characteristics we control for sample heterogeneity. In particular we show that the introduction of the unobserved heterogeneity term is important, but the assumption about its functional form is not. The case of a non parametric specification of the random effect identifies two different groups. The first can be though about as "movers", i.e. women considered more family oriented and making the transition faster than the "stayers".

In order to asses which wage level captures the largest part of the variance of the unobserved heterogeneity, we interact the random effect with different socio-economic groups (i.e. predicted wages). Women with relatively low and high wages are the typology of women for which it is more difficult to explain their attitude towards fertility decisions without considering other unobserved characteristics.

Chapter 3 tackles a more methodological issue. It presents a series of Monte Carlo simulation-based analysis to study the consequences of the omission or misspecification of unobserved heterogeneity in single spell discrete time duration models. A preliminary version of this chapter has been published as a working paper (n. 2006-53) of the Institute for Social and Economic Research and it is a joint work with Cheti Nicoletti from ISER.

Empirical researchers usually use statistical models which can be easily estimated with the help of commonly available software packages. Sequential binary models with or without normal random effects are an example of such models, because they can be adopted to estimate discrete time duration models in presence of unobserved heterogeneity. But an easy-to-implement estimation may incur a cost. We assess the effects of ignoring Unobserved Heterogeneity using time variant, time invariant and mixture explanatory variables and of misspecifying its distribution imposing a Gaussian Random Effect when the real distribution is a gamma or discrete one with a specific emphasis on discrete time hazard models.

We extend the well known results for continuous time duration models to discrete one. In particular for continuous models, Ridder (1987) using Monte Carlo Simulations and Dolton and van der Klaauw (1995), Meyer (1990) and Trussell and Richards (1985), empirically, find that omitting the Unobserved Heterogeneity, if the duration dependence is modeled in a flexible way, do not bias the estimated coefficients.

For discrete time single spell models, Baker and Melino (2000) find that even when controlling for a flexible duration, but ignoring the Unobserved Heterogeneity, the parameters are biased. We believe this result is due to the fact that sequential binary models are identified up to a scale normalization and models with different specification may imply a different normalization.

We find that ignoring the unobserved heterogeneity in sequential logit models causes an overestimation of the negative duration dependence and an attenuation of the covariate coefficients but this attenuation is due to the fact that all the coefficients are rescaled by the same factor. Since coefficients in binary models are always identified up to a scale normalization, inference should not be affected by the unobserved heterogeneity omission except for the duration dependence. Furthermore, it seems that misspecification of the unobserved heterogeneity distribution does not affect the estimation results.

Changes in the error distribution (logistic, Gaussian and extreme value) bias the duration dependence estimation but cause only a rescaling by a factor of the estimated coefficients. Given that coefficients in binary models are identified only up to a scale normalization, the rescaling by a factor issue is not problematic.

Chapter 1

Women's Wages and Childbearing Decisions: Evidence from Italy.

1.1 Introduction

During the early 1990s, Italy became one of the first countries to reach lowest-low fertility, i.e. a Total Fertility Rate (TFR) below 1.3. Whereas the reasons behind the Italian fertility decline are many, a prominent explanation suggests that it has been driven by increasing opportunity costs. Certainly, the strong fertility decline during the nineties was followed by a significant increase of education and labour force participation among women. Given this setting it is of interest to establish the extent to which women's wages are linked to their fertility decisions. As a result we analyse fertility behavior of Italian women during the period 1983 to 2003. We pay particular attention to the role of potential incomes (or predicted wages), which in turn depend directly on women's educational attainment, and therefore their opportunity cost, and we investigate how this affects their childbearing decisions.

So far research on this issue in Italy has been hampered by lack of data sources that contains information on both fertility behavior and women's income or earnings. We overcome this problem by combining two different data sets. Data from the Italian Institute of Statistics Labor Force Survey of 2003 (ISTAT-LFS from now on) are used to reconstruct women's fertility histories. Potential wages (i.e. hourly labour income) are derived from the income and earnings data from the Bank of Italy Survey of Households' Income and Wealth (BOI-SHIW, 2002 from now on). Data from the BOI-SHIW are thus used to predict women's potential wage, which is then introduced as an explanatory variable in discrete time hazard regression models for first, second, and third births.

Our analysis includes women aged between 15 and 40 years in 2003. Their predicted wages depends on a range of background variables, including their educational attainment. As such the predicted wages reflect their potential wage, which in turn is a measure of their socio-economic characteristics. Thus our interest lies in assessing the extent to which socio-economic characteristics may play a key role in the postponement of motherhood in Italy and its differential impact in the transition to first, second and third birth. In particular, we assess whether differences in socio-economic characteristics are the only determinants in these transitions, or whether other socio-cultural features might be responsible for delayed motherhood and its impact on second and third birth decisions.

The remainder of this chapter is organized as follows. In Section 1.2 we review the existing literature and describe the main features of the Italian setting as far as childbearing decisions are concerned. Section 1.3 provides a description of the data and methods we use. In section 1.4 we present our main results. Section 1.5 includes concluding remarks and some policy considerations.

1.2 Background

1.2.1 Previous Literature

It is useful to start from an economic perspective when considering the role of wages and income on fertility. Becker (1960) argues that there should be a positive correlation between the number of children and the household income. Highly educated mothers, however, tend to substitute the number of children with child quality (see Becker and Lewis 1973). Since both production and rearing of children are time intensive, an increase in wage rates induce a negative substitution effect on the demand for children (see for instance Becker 1965; Mincer 1963). To this extent, theoretical research on fertility (like Hotz, Klerman and Willis 1997; Becker 1981; Willis 1973) shows that women's income is negatively associated with childbearing as a higher income implies higher opportunity cost of children. In other words, with high earnings, it becomes more expensive for her to take time away from work to rear children. In general the opposing income effect is unlikely to outweigh the negative substitution effect. Overall therefore we would expect the effect from women's wages to be negative. For men, in contrast, the income effect tends to dominate since they spend less time on rearing children, though the magnitude of these effects will vary across countries and birth parity (Butz and Ward 1979; Willis 1973).

Ermisch (1989) provides an important extension to the Beckerian argument - the insight being that women's fertility decision does not only depend on their wages, but also on availability of external child-care. In the presence of costly external child-care, women with very high earnings, traditionally having a high negative opportunity cost of childbearing, may instead have more children, because they are more able to afford external child-care.

Those with very low wages are less likely to afford external child-care, but may still have higher fertility due to low opportunity costs. In this scenario both low and high income women will have more children, whereas those with middle income will have lower fertility (i.e. lower demand for children). The argument depends of course on the availability of child-care. We might expect such effects in Scandinavian countries, whereas in Italy we expect to see more of the traditional pattern.

In recent years research has shifted towards investigating the timing of births rather than completed fertility (see for instance Heckman and Walker 1990a). In these studies, hazard rate models are used to empirically analyse the timing and spacing of births. Most of the studies show that, compared to women with low opportunity costs, women with high wages (i.e. high opportunity cost of having children) have births later. There is also ample evidence to suggest that presence of children has a significant negative impact on the woman employment probability (see Mroz 1987; Heckman and MaCurdy 1980). Thus, women's labour supply and fertility are joint decisions, and cannot be considered as separate processes.

Likewise, the decision of entering parenthood has important time dimensions to it. A couple may want to pay attention to the cost of having children *early*, comparing it with the cost of having them *later*. In such a way, the optimal age of having the first birth can be seen as a trade-off between investment in human capital and career planning (see Gustafsson 2001). As emphasized by Cigno and Ermisch (1989), Cigno (1991, chapter 8) and Gustafsson and Wetzels (2000), it is important to take into account possible consequences of lifetime earnings given different scenarios of birth timing and spacing. In light of this it may be optimal for many women to delay motherhood until the opportunity cost of child-care (with respect of her career) have decreased, which implies completion of education and getting a foothold in the labour market, before entering motherhood.

Formally, timing of first birth is a function of the opportunity cost of time and the foregone human capital cost (Gustafsson 2001; Cigno, 1991). However, the effect of income on the timing of births may work through different paths. Gustafsson (2005) suggests that, for Swedish young individuals, the postponement of formal education works by delaying couple formation rather than by delaying parenthood once the couple is formed. So, family policies may have a pro-natalist effect in allowing Swedish couples to have the first child earlier. A somewhat different interpretation is offered by Amuedo-Dorantes and Kimmel (2005) asserting that college-educated mothers can profit from delaying motherhood, because they are in a position to negotiate a family-friendly work environment with flexible work schedules.

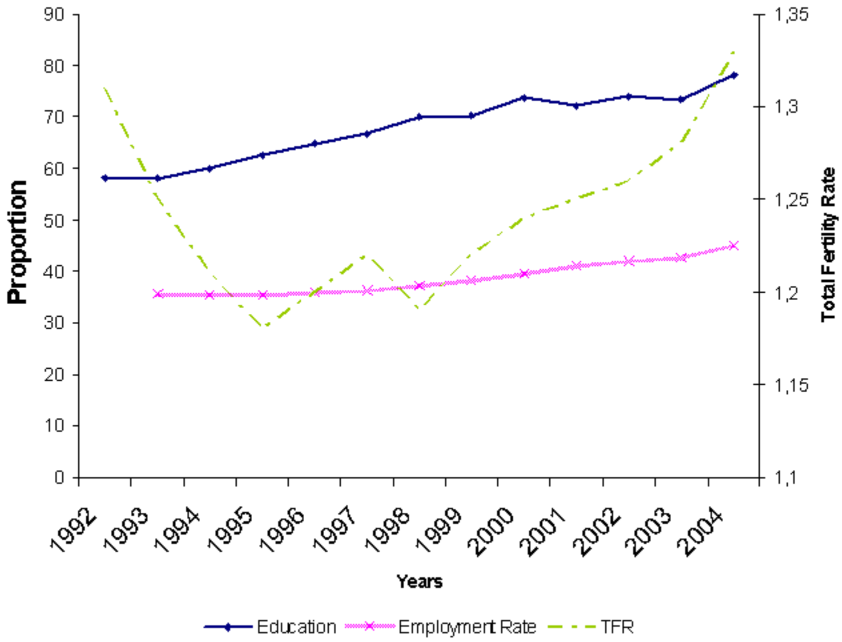
An important part of the existing literature argues that women's responsibility for child-rearing may reduce her time in paid work. Joshi (1990), for instance, analyse how work patterns can be different (in terms of switching from full time to part time work or not employed) by comparing mothers with childless women. The causal direction is explained through the impact of children on the women's job market opportunities and their level of income. Miller (2005) and Waldfogel (1998) offer similar perspectives.

1.2.2 The Italian setting

Figure 1.1 provides stylized facts of fertility in Italy.

Two important trends are prevalent for the last two decades in Italy: a sharp decline in total fertility and a steady increase in women's educational attainment, which is followed by higher female employment rates (see also Figure 1.2). For that period and using a sensitivity analysis, Rindfuss, Guzzo and Morgan (2003) showed that a 1% increase in female labour force participation was associated with a 3.15% decline in Total Fertility Rate. During the same period, the mean age of onset of motherhood increased from 25 years in 1980 to almost 27 in 1990, reaching a level of 28 years in 1995 and 28.7 in 1997 (Figure 1.2). As Kohler, Billari and Ortega (2002) point out, Italy has been, together with Spain, the first country to reach the threshold of so-called lowest-low fertility, i.e. below 1.3 children per woman.

Figure 1.1: Education, Fertility and Woman Participation Rates in Italy (1992-2004).

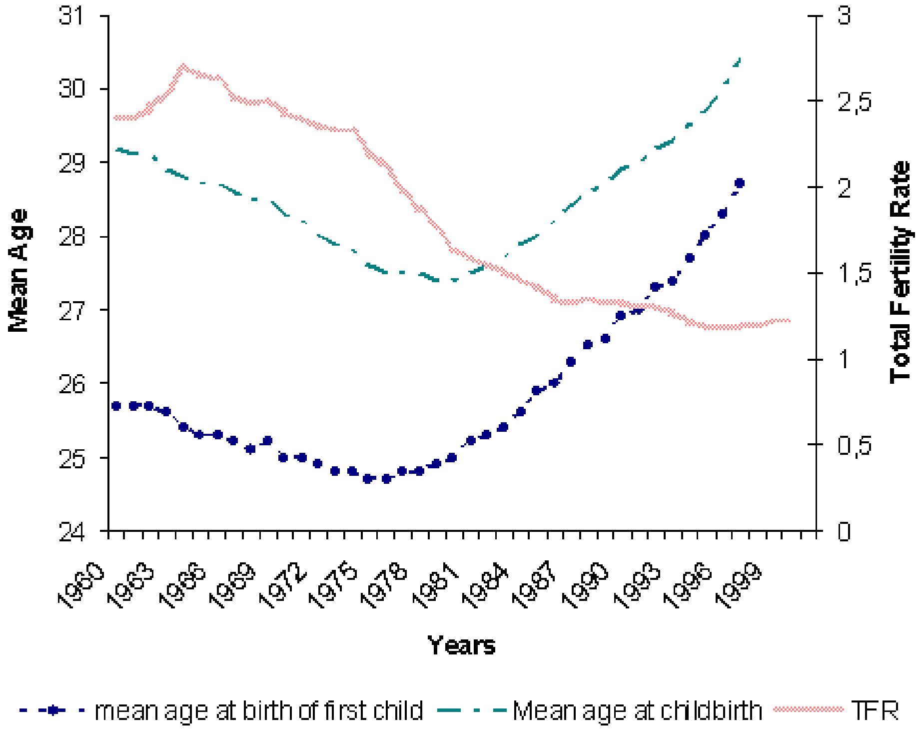


Notes: Fertility rates: 1998 and 1999 are provisional values and 2000, 2001, 2002 are estimated values. Education is the percentage of the female population aged 20 to 24 having completed at least upper secondary education. The female employment rate is calculated by dividing the number of women aged 15 to 64 in employment by the total female population of the same age group.

Source: EUROSTAT. http://epp.eurostat.ec.europa.eu.

It is well established that the emergence of the lowest-low fertility in Southern Europe is not connected to any steep increase in childlessness (Billari and Kohler 2004). Available parity-specific data on fertility show that most of the fertility decline in Italy during the last twenty years is due to the sudden decrease of the progression to the second, third and subsequent children. As a consequence, the probability to have a first child has not changed in spite of the tremendous economic and social changes characterizing Italy during the second half of the 20th century (see Dalla Zuanna 2004). Moreover, the personal ideal family size for around 60% of Italian women aged 20-34 years is two children; while one quarter have a preference for large families (Goldstein, Lutz and Testa 2003).

Figure 1.2: Mean age at first birth, at childbirth and TFR in Italy (1960-2000).



Notes: Mean age of women at childbearing is the mean age of women when their children are born. For a given calendar year, the mean age of women at childbearing can be calculated using the fertility rates by age (in general, the reproductive period is between 15 and 49 years of age).

Source: Council of Europe (2001), *Recent Demographic Developments in Europe.* http://www.coe.int/t/e/social_cohesion/population/demographic_year_book/2001_edition/Italy%202001.asp#TopOfPage.

Economic factors might thus shape a) the timing of first birth more than the probability of ever having a first birth and b) the timing and the probability of ever having births of higher order. Data from opinion surveys support this idea. In a recent survey (2002) on a sample of mothers aged around 42, living in five Italian cities, women gave reasons for why they had stopped at the parity they actually experienced. Concerning the transition to the third birth (i.e. women who stopped at two children), economic reasons were cited as important for women who experienced a worsening of their financial situation after the birth of the first and second child. Many women argued that monetary transfers for the first three years after the birth of a third child, or a lower but longer financial incentive, could have changed their decision to stop at parity two (De Santis and Breschi 2003). Although being possibly biased as an ex-post motivation, this role of economic factors is specific for third birth.

It is thus reasonable to suppose that increasing financial support to households with children may have an impact on the probability of having the third and fourth child, mainly for poorer households (De Santis and Breschi 2003). To this extent, the Welfare Minister Livia Turco (law number 448 of the Year 1998, Turco's law from now on) introduced two policy measures with the explicit purpose of supporting poor households with children. The measures of the law were introduced in 1999. The first measure provided a cash transfer of around 110 Euros per month for households with at least three children under 18 who had low household income levels (i.e. less than 15,000 Euros a year before taxation). This amount grew slowly year on year, following the life-cost index and it reached around 120 Euros a month in 2001. The share of households receiving this transfer has been particularly sizeable for larger households, especially in Southern Italy (Lelleri and Marzano 2002). About 300 million Euros were transferred in total in 1999 and 2000. The second measure had relatively mild restriction on income levels and it provided a monetary transfer to households in which one of the partners (typically woman) was not employed. The transfer, for a period of 5 months, was a monthly amount of 100 Euros in 1999, 155 in 2000, 260 Euros in 2001 and 2002. A significant share of women received this transfer, especially in the South (Lelleri and Marzano 2002). The two measures could also be simultaneously received, in an additive way. They were introduced for the anti-poverty purpose of assisting families with many children, who are at risk of being poor; they were not introduced as pro-natalist measures, but could be implicitly pronatalist (Whittington, Alm and Peters 1990). Indeed, these two measures could cause a significant increase in income for low-income households, covering a non-negligible proportion of the cost of an additional child.

1.3 Data and Methods

Micro evidence on the relationship between the timing and the spacing of births has been scarce for the Italian case, the main reason being a lack of data sources that consist of wage and income information that can be linked with women's fertility histories and where the latter has sufficient sample size. The data requirement for estimating the impact of wage on the timing of births is demanding: we need panel data with a long time dimension or retrospective data on the complete employment and fertility histories. Usually, and this is our case, a researcher has available only cross section data or panel data with a short time dimension. Fertility histories may be reconstructed on the basis of the age of the children in the household but the complete labour market history of the women in the household will be more difficult or even impossible to reconstruct. This is because one observes for each household the number of children present and the employment status of the women in the household only at the time of the interview.

Unfortunately none of the currently available Italian data sets contain all the required information. The BOI-SHIW, 2002, contains detailed information on employment and income of family members, labour market activities, payment instruments and forms of savings, socio-demographic characteristics of the household. However, the sample size is too small to conduct fertility analysis, particularly for third births. The ISTAT-LFS, 2003, provides detailed information on the family structure, labour market, work experience, part time and full time employment. The main drawback of this survey is that it does not collect information on household earnings and income. The sample size is however large and suitable for fertility analysis.

In order to overcome these limitations, we combine the two data sets. We use the BOI-SHIW data set to estimate earnings equations, from which the predicted earnings are matched onto women in the ISTAT-LFS data set.

1.3.1 Bank of Italy Survey of Italian Households' Income and Wealth (BOI-SHIW)

The BOI-SHIW started in 1964. Its main aim was to collect information on income and savings of Italian households. During the past few years, however, new questions about payment instruments and different forms of savings were introduced. The survey collects information on more than 22,100 individuals (8,011 households) with 13,536 individuals receiving an income. From the survey we have information about the activity of the employees (their total net income, average worked number of hours, hours of paid overtime), their professions, whether they are sole proprietors or free-lances, contingent worker employed on none account (if they worked all year or only for part of the year, their net earnings and average worked number of hours), about the family businesses, active shareholder/partner, pensioners and other income such as scholarship, alimony etc. For each individual we have the annual income and their wages.

Our unit of analysis is the woman. We only consider households where it is possible to link every woman with their co-residing children. In total this produces a sample of 20,003 individuals with 4,749 women in their childbearing stage. In other words, we consider women aged less than 45 years, who are not pensioners or not studying.

1.3.2 The Italian Institute of Statistics Labor Force Survey (ISTAT-LFS)

The ISTAT-LFS is a quarterly and continuous survey implemented since 1959. For each year four waves are carried out. The survey collects information on more than 300,000 households, which constitutes around 800,000 individuals (1.4% of the total national population) distributed over 1,351 municipalities (out of 8,000). The ISTAT-LFS is the principal data source for assessing the Italian labour market. The sample design is a two stage rotating sample design with stratification of the primary units (municipalities). Each household is included first in two waves, then left out for two waves and then included in another two waves. The ISTAT-LFS offers different sections dealing with demographic characteristics of the households, present job (with all the information taken from the month before the interview), job experience, looking for a job, relationship with public employment centres. However, there are no retrospective fertility histories available. Instead we know the number of children (and their age) living in the households where the woman is either the household head or the spouse of the household head.

Fertility histories are derived from recorded co-residing children at the time of interview. We are able to reconstruct retrospective fertility histories for each women back to 1983. The sample includes a small amount of adopted children or stepchildren but we exclude any offspring who might have died or moved away. In Italy mortality at adult age is low and children of divorced parents are almost exclusively living with their mothers and a very low proportion of young individuals leave the parental household before 23 (Aassve, Billari, Mazzucco and Ongaro 2002). Nevertheless, in order to ensure that the recorded children are the only ones of the mother, we limit the analysis to only include women who are aged 40 or less in 2003. Given that the mean age of leaving home is rather high in Italy (Aassve, Billari, Mazzucco and Ongaro 2002) this selection does not reduce the sample by much. We also exclude households where we were unable to link children with mothers (i.e. male head of the household with no wife and all single men). In this way we have a sample of 34,914 women. In total, the fraction of women that can be matched with their co-residing children is 95%. Unfortunately, we are not able to use information of husbands since we only know the marital status of women at the time of interview. As a result we are prevented from reconstructing retrospective marriage histories. This problem also applies to widows in that we do not know when the husband died.

Table 1.1 and Figure 1.3 report the relevant sample statistics. Given our sample of women aged 15 to 40, Table 1.1 shows that the mean age at first birth is 26.2 years when evaluated

Table 1.1: Mean age at first, second and third birth per Regions and level of Wages.

		Wage $\leq$ 25th percentile	Wage $\geq$ 75th percentile
Transition to 1st birth:			
	26.2	22.2	29.6
South	25.0	22.5	30.5
Centre	26.7	20.5	29.7
North	27.1	18.6	29.4
Transition to 2nd birth:			
	28.9	25.9	32.1
South	27.9	26.2	32.7
Centre	29.5	25.1	31.9
North	29.9	21.9	32.0
Transition to 3rd birth:			
	30.8	29.2	33.4
South	30.4	29.4	34.1
Centre	31.3	27.7	33.1
North	31.3	23.8	33.2

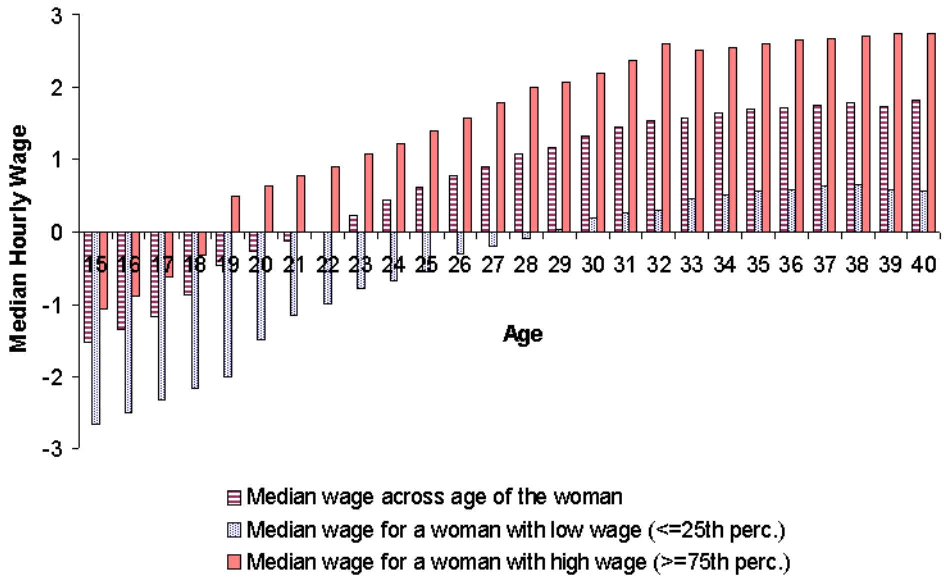
Notes: The percentiles refer to the distribution of predicted hourly wage from the estimated wage equation as reported in Appendix A. The means are for closed birth interval only (computed only for women experiencing a birth). The 20 Italian Regions are classified according to the following. Piemonte, Valle d'Aosta, Lombardia, Liguria, Trentino, Veneto, Friuli Venezia Giulia, Emilia Romagna belong to the North. Toscana, Umbria, Marche, Lazio are considered Central Regions. Abruzzi, Molise, Campania, Puglia, Basilicata, Calabria, Sicilia, Sardegna belong to the South.

Source: Own Calculation from ISTAT-LFS, 2003, with fertility histories retrospectively reconstructed from 2003 back in 1983.

over the sample period from 1983 to 2003. Because of the non homogenous institutional background, the regions are stratified into North, Center and South. While poor women in the North (where poor means an hourly wage lower than the 25th percentile of the wage distribution for the whole country) starts to have children early, women in the South tend to delay motherhood until they are 22.5. Transition to second birth takes place on average three years after the first child is born with high wage mothers in the South delaying more than the other ones. The third birth, instead, seems to succeed the second one with a shorter interval: if a woman is having a third birth she only wait, on average, 2 years since the previous one. Again, low wage mothers living in the North tend to concentrate their fertility history in a shorter time span: they start at 18.6 and end it after 5 years (conditional on reaching the threshold of the 3rd birth).

Figure 1.3 reports the financial situation of an average woman, where average refers to mean centered wage. Median wage increases over time at decreasing rate. For a low wage woman, it grows quickly until it reaches the zero threshold at age 29, but a different setting is offered by a high wage woman who shows a positive, albeit slow, increasing wage since the age of 19.

Figure 1.3: Median Hourly Wage per age of an average woman.



Source: Own Calcutation from ISTAT-LFS, 2003 where wage is the predicted hourly wage estimated in Appendix A. A negative wage means it is smaller than the average wage per age.

1.3.3 Methods

Our methodological strategy can be summarized in three steps. First, we estimate wage equations based on the BOI-SHIW data. Second, we match predicted wages onto the ISTAT-LFS data using common characteristics of women in both data sets. Third, we create three sub-samples consisting of women being at risk of the first, second and third birth.

The details of estimation of wage equations is given in Appendix A. The estimation of the impact of predicted (and time-varying) wage $\hat{\omega}_{it}$ on fertility is implemented through a set of discrete time event models. Consider a series of P predictors $X_{1ij}, X_{2ij}, ..., X_{Pij}$ and let x_{pij} denote individual i 's values for the p-th predictor in time j. The hazard function is defined as:

$$h(t_{ij}) = Pr[T_i = j | T_i \geq j \,\text{and}\, X_{1ij} = x_{1ij}, X_{2ij} = x_{2ij}, ..., X_{Pij} = x_{Pij}] \tag{1.3.1}$$

that is, the population value of discrete-time hazard for person i in time period j is the probability that he/she will experience the target event in that time period, *conditional* on no prior event occurrence *and* his or her particular values for the P predictors in that time period (Jenkins 1995).

We estimate a variety of models and the specified baseline is different depending on birth parity:

A. *Baseline for the model of the first birth* is:

$$logit(h_{ij}) = [\alpha_{15} A_{i15} + ... + \alpha_{40} A_{i40}] \tag{1.3.2}$$

where A_{ij}, $j = 15, ..., 40$ are dummy variables indicating a non-parametric specification of time and

$$A_{ij} = \begin{cases} 1 & \text{if } i\text{-th woman is } j\text{-years old} \\ 0 & \text{Otherwise} \end{cases} \tag{1.3.3}$$

so that the baseline for the model of the transition to first birth is a function of the age of the woman.

B. *Baseline for the model of the second birth* is:

$$logit(h_{ij}) = [\beta_1 D_{i1} + ... + \beta_{25} D_{i25}] \tag{1.3.4}$$

where D_{ij}, $j = 1, 2, ..., 25$

$$D_{ij} = \begin{cases} 1 & \text{if } i\text{-th woman is observed after } j\text{-years from the birth of the first child} \\ 0 & \text{Otherwise} \end{cases} \tag{1.3.5}$$

so that D_{ij} is a function of the spell, duration, from the first birth. (The baseline is similarly specified for the third birth).

In order to asses the impact of wage of fertility different specification of the model are offered:

C. *General Wage Effect*:

$$logit(h_{ij}) = [\delta_1 P_{i1} + ... + \delta_J P_{iJ}] + \pi_1 \hat{\omega}_{ij} + \pi_2 \hat{\omega}_{ij}^2 \tag{1.3.6}$$

where P_{ij}, $j = 1, ..., J$

$$P_{ij} = \begin{cases} A_{ij} & \text{if first birth is under study} \\ D_{ij} & \text{if second/third birth is under study,} \end{cases} \tag{1.3.7}$$

A_{ij} and D_{ij} specified as in (1.3.3) and (1.3.5) respectively and $\hat{\omega}_{ij}$ is a time-varying covariate, predicted according to equation (A.0.5) as showed in the Appendix A.

D. *Wage Effect and Duration Effect*:

$$logit(h_{ij}) = [\delta_1 P_{i1} + ... + \delta_J P_{iJ}] + \pi_1(\hat{\omega}_{ij} * P_{i1}) + \pi_2(\hat{\omega}_{ij} * P_{i2}) + ... + \pi_J(\hat{\omega}_{ij} * P_{iJ}) + \gamma H_i \tag{1.3.8}$$

where

$$H_i = \begin{cases} 0 & \text{if first birth is under study} \\ \text{age first child}_i & \text{if } i\text{-th woman is at risk of second birth} \\ \text{age second child}_i & \text{if } i\text{-th woman is at risk of third child,} \end{cases} \tag{1.3.9}$$

P_{ij}, $j = 1, ..., J$ is specified as (1.3.7) indicating the duration effect, $\hat{\omega}_{ij}$ is a time-varying covariate for the i-th woman at time j, predicted according to equation (A.0.5), age first child$_i$ is the age at which the i-th woman had the first child (a woman is at risk of the second child since she had the first one), age second child$_i$ is the age at which the i-th woman had the second child and the term $\hat{\omega}_{ij} * P_{ij}$ simultaneously captures the wage and the duration effect for the i-th woman at time j.

E. *Institutional/Cultural Effect*:

$$\begin{aligned} logit(h_{ij}) &= [\delta_1 P_{i1} + ... + \delta_J P_{iJ}] + \xi_k(\hat{\omega}_{ij} * \sum_{k \in \{S,N,C\}} d_{ik}) + \tau_k(\sum_{k \in \{S,N,C\}} d_{ik}) + \\ &+ \gamma H_i + \pi_1(\hat{\omega}_{ij} * P_{i1}) + \pi_2(\hat{\omega}_{ij} * P_{i2}) + ... + \pi_J(\hat{\omega}_{ij} * P_{iJ}) \end{aligned} \tag{1.3.10}$$

where $k \in \{South, Centre, North\}$,

$$d_{ik} = \begin{cases} 1 & \text{if } i\text{-th woman lives in the } k\text{-th geographical part of Italy} \\ 0 & \text{Otherwise} \end{cases} \tag{1.3.11}$$

P_{ij} specified as in (1.3.7), $\hat{\omega}_{ij}$ estimated according to equation (A.0.5), H_i defined in (1.3.9) and $\hat{\omega}_{ij} * P_{ij}$ simultaneously captures the wage and the duration effect for the i-th woman at time j.

F. *Turco's Law (1999)*:

$$logit(h_{ij}) = \tau_k(\sum_{k \in \{S,N,C\}} d_{ik}) + N_i + F_i + F_i * N_i \tag{1.3.12}$$

where $j = 1983, ..., 2003$,

$$F_i = \begin{cases} 1 & \text{if } i\text{-th woman has low wage (predicted distribution)} \\ 0 & \text{Otherwise,} \end{cases} \tag{1.3.13}$$

d_{ik} defined in (1.3.11) and

$$N_i = \begin{cases} 1 & \text{if } j \geq 1999 \\ 0 & \text{if } j < 1999. \end{cases} \tag{1.3.14}$$

In all our discrete time duration models (estimated from ISTAT-LFS), the wages are predicted from BOI-SHIW. The standard errors are consequently corrected for the fact that we use generated regressors in the fertility equation (see Pagan 1984; Arellano and Meghir 1992). To do so we use the bootstrap method with 300 replications and observations clustered at the woman level. Any further increase in replications does not significantly affect the estimated standard errors.

1.4 Results

In this section we present results showing how predicted wages may affect the risk of having the first, second and third birth. We conclude the section by including the empirical evidence concerning Turco's Law (1999). The estimations ignore the effect of the man's wage, because we are not able to go back and reconstruct retrospectively when a woman got married (we only have information about the marital status in 2003).

1.4.1 First Birth

As expected the baseline hazard for first births depicts an inverse U-shape. The effect of wages is approximated by a second order polynomial, which simply means that we include wages and its squared term in the regression. Column (1) of Table 1.2 shows that the general effect of women's wage is negative: the higher the wage, the lower is the risk of entering motherhood (-0.407). The effect of wage is however nonlinear which is confirmed by the positive coefficient associated with the square term. If controlling for wage and its squared term only, the mean age of first birth is estimated to be between 30 and 31 years of age. This effect is however an overestimate. In 1997, for example, the average age at first birth was 28.7 (See Council of Europe (2001), *Recent Demographic Developments in Europe*). As shown in Figure 1.4, once we control for interactions between age and wage, the baseline hazard is shifted to the left (an anticipation effect), resulting in a correct mean age for the first birth being between 27 to 28.

In column (3) and (5) of Table 1.2 we report the estimated coefficients for women being at risk of the first birth where wage has been stratified by age and interacted with region. This

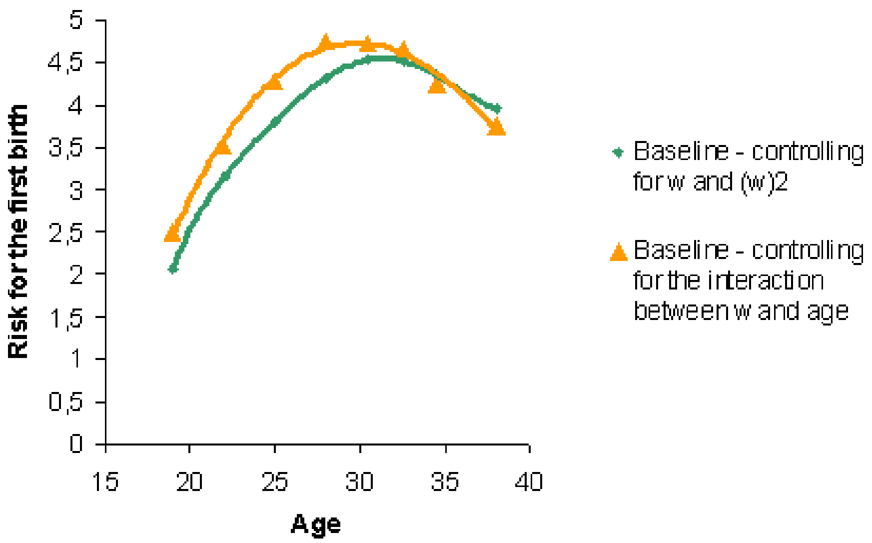
Table 1.2: Transition to First Birth and Wage Effect

Variable	Coeff.	(Std. Err.)	Coeff.	(Std. Err.)	Coeff.	(Std. Err.)
	(1)	(2)	(3)	(4)	(5)	(6)
General Wage Effect:						
Wage	-0.407**	(0.007)				
$(\text{Wage})^2$	0.070**	(0.003)				
Wage and age:						
Wage × age (15-17)			-0.770**	(0.074)	-0.698**	(0.076)
Wage × age (18-20)			-0.508**	(0.020)	-0.436**	(0.021)
Wage × age (21-23)			-0.620**	(0.016)	-0.540**	(0.017)
Wage × age (24-26)			-0.494**	(0.013)	-0.409**	(0.015)
Wage × age (27-29)			-0.330**	(0.014)	-0.244**	(0.017)
Wage × age (30-31)			-0.104**	((0.021)	-0.021	(0.024)
Wage × age (32-33)			-0.064*	(0.025)	0.018	(0.028)
Wage × age (34-35)			0.070†	(0.039)	0.152**	(0.041)
Wage × age (36-40)			0.121*	(0.052)	0.206**	(0.054)
Region effect:						
Center					-0.244**	(0.023)
North					-0.191**	(0.020)
Wage and region:						
Wage× Center					-0.008	(0.019)
Wage× North					-0.058**	(0.016)
Constant	-6.167	(0.056)	-6.561**	(0.157)	-6.355**	(0.163)
No. of Observations	425,014		425,014		425,014	
No. of women	34,129		34,129		34,129	
No. of Spells	25,281		25,281		25,281	
Pseudo R^2	0.084		0.087		0.088	
Log-L	-87643.313		-87354.001		-87265.406	

Notes: Discrete-time logit hazard regressions. Baseline not reported. Reference category for Region is South. Columns (1), (3) and (5) correspond to equations (1.3.6), (1.3.8) and (1.3.10) respectively. Bootstrapped standard errors in parentheses. Standard errors are computed using 300 bootstrap samples.

$^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$.

Figure 1.4: Baseline for the First Birth

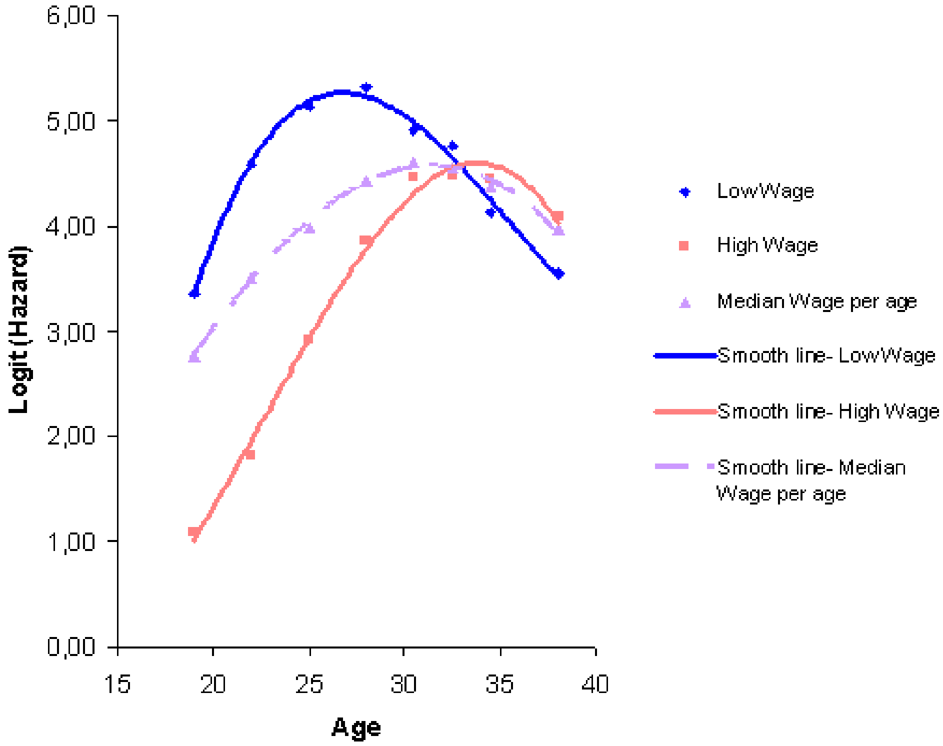


Notes: Estimated baseline from equation (1.3.6) and (1.3.8). Women between 15 and 18 are the reference group.

allows us to understand how wage impacts on first birth after controlling for age and cultural or institutional effects. Column (3) of Table 1.2 shows that wage has a strong negative effect on the risk of the first birth for young women, the effect becoming closer to zero at the age of 34-35. The result shows that women with high wages have a higher likelihood of postponing motherhood, and confirms the importance of wages in driving the onset of motherhood. When controlling for regions (column (5) of Table 1.2), we mostly find the same negative pattern for younger women but the turning point is now between 32-33 years. After 34 years, increasing wage by one unit increases the likelihood to experience motherhood. Women in Central Italy are less at risk of first birth when compared with women in the South: a similar pattern (bigger in absolute value) is showed by Northern mothers (column (5) of Table 1.2). However, when controlling for cultural and institutional effects (Wage and Region effect of

Table 1.2), high wage Northern women are less likely to become mothers for the first time compared to the Southern one.

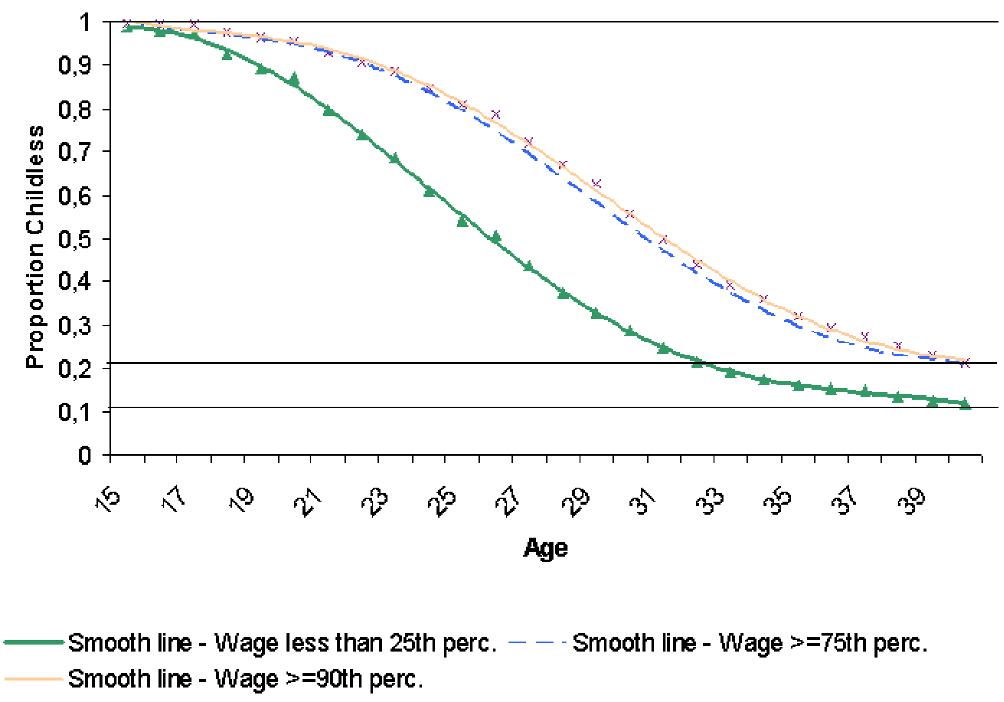
Figure 1.5: Timing of first birth for a low and high wage woman



Notes: Logit(Hazard) estimated from Table 1.2, column (3). Low and High wage are constant through age: low (high) wage are set equal to the 10th (90th) percentile of the hourly wage distribution.

In order to better understand how different wage levels affect the timing of first birth, we simulate the hazard paths for a low and high wage woman. We choose two extreme situations. A low wage woman is assigned a wage set to the 10th percentile of the hourly wage distribution, while a high woman's predicted wage is set to the 90th percentile. Both paths are plotted together with the hazard of the median wage woman in Figure 1.5. The Figure shows a non proportional hazard: low wage women are at a higher risk of experiencing the first birth when they are very young, reaching the maximum level when they are aged between 25 and 30 years of age. High wage women tend to delay, maximizing the likelihood for the first birth when they are 30 or older. The median wage woman has, not unexpectedly, an intermediate position. When the three paths cross (at age 32) low, median and high wage women experiences similar risks though with the difference that low wage women have already reached their maximum, the median woman is at the maximum risk level and the high wage woman is yet to reach the maximum.

Figure 1.6: Predicted Survival Curves. Postponement and Recuperation Effect.



Notes: Median predicted survival per ages with wage evaluated with a mobile threshold (25th, 75th, 90th percentile) per ages. The proportion of women childless is estimated from Table 1.2, column (3). Recuperation effect is the vertical distance between the three predicted survival curves.

In addition to plotting the hazard rates, we also plot the simulated survival curves in Figure 1.6. A highly interesting pattern here is the strong recuperation effect among high wage women. Though the recuperation is not complete, the Figure suggest that high wage women are in effect quite similar to low wage women in terms of completed fertility of first parity. In other words, high wage women start the onset of motherhood later, but recuperate after some time. This pattern is rather stable, and persists also when the high wage women are defined according to the 75th percentile as opposed to the 90th percentile. The same is true if low wage women are defined according to the 25th percentile. We find that higher wages have their primary effect on postponement of first birth so that when wages are higher, pregnancy tends to be concentrated in a shorter span of the life cycle that starts later in life. This is in line with the opportunity cost theory (see Heckman and Walker 1990a): few employed mothers would quit their job to have children if an exit from the labour market could seriously damage their future labour market prospects (Boix 1997). This effect would be more important the greater the uncertainty in the labour market or the higher the unemployment rate, and is one of the reasons why Central and Northern regions exhibits negative estimated coefficients for the risk of the first birth when compared to the South (where there is higher labour market uncertainty).

The paths may also be explained in terms of recruitment policies which differ across regions. As many firms often recruit young talents prior to motherhood rather than women in the post-parental phase, this might reinforce postponement incentives (Gustafsson 2001). The negative coefficients of the interaction between wage and region (column (5) of Table 1.2) can therefore be explained by the large presence of public jobs in the South where working women are mostly employed in teaching or administrative activities, mainly driven by a lack of private firms. Women working in public sectors tend to have permanent contracts and pre/post maternal paid periods, providing greater stability (i.e. low likelihood experiencing job redundancy). In the North, in contrast, the style of work life changes considerably, with a much higher prevalence of private firms. To this extent women from the North, normally with higher wage levels, are forced to postpone motherhood until they reach a certain position in their career planning. Consequently depreciation of women's human capital due to temporary absence from the labour force due to childbearing, might be more severe compared to women in the South (Bernardi 2003 offers a qualitative research on the role that social influence plays on the transition to motherhood in Italy).

Another argument which is consistent with the negative interaction between wage and region is that, presumably, mothers with higher career prospects decide to postpone maternity until they obtain a more stable labour market situation (e.g. getting a permanent contract (see de la Rica and Iza 2005 for an explanation for the Spanish case). Unfortunately, our data reconstruction of women's job history makes it difficult to verify these hypothesis. However, it seems there is no institutional/cultural effect driving the first birth decision. If this was the case, we should have found a positive coefficient for the interaction of wage with North where a more efficient child care system is set up in order to help working mothers to conciliate job and family decisions.

1.4.2 Second and Third Birth

A woman becomes at risk of second birth once she experienced the first. Similarly, she becomes at risk of the third once she has experienced the second. A trivial consequence of this obvious fact (yet important for completed fertility) is that the age of a woman at the time she becomes at risk may affect second (and third) birth fertility (see Heckman and Walker 1990b). Our estimates in Table 1.3 prove that the higher the age of first birth, the less is the risk of second birth (-0.014).

Table 1.3: General Wage Effect for Second and Third birth.

Variable	**Coefficient**	**(Standard Error)**	**Coefficient**	**(Standard Error)**
	Transition to second birth		Transition to third birth	
General Wage Effect:				
Wage	-0.073**	(0.009)	-0.0004	(0.022)
$(\text{Wage})^2$	0.066**	(0.005)	0.099**	(0.012)
Age at first birth	-0.014**	(0.003)		
Age at second birth			-0.102**	(0.007)
Constant	-3.615**	(0.078)	-2.166**	(0.207)
No. of Observations	147,539		98,418	
No. of Women	25,227		14,483	
No. of Spells	14,334		2,506	
Pseudo R^2	0.086		0.047	
Log-L	-42928.736		-11103.104	

Notes: Discrete-time logit hazard regressions for second and third birth. Baseline not reported. Estimated equation (1.3.6). The number of women at risk of a second birth differs from the number of spells reported in Table 1.2 because women having twins as first births are not in the risk set of second birth, but they are directly at risk of a third birth. Bootstrapped standard errors in parentheses. Standard errors are computed using 300 bootstrap samples.

$^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$.

The coefficient for the age of second birth is smaller than the previous one, meaning that the risk of having a third child is strongly negatively correlated with the age at which second birth took place. In other words, mothers who tend to delay first birth are at a lower risk of having a second birth and those who delay the second have lower likelihood of having a third. Also in this case wage has a negative effect on fertility at a decreasing rate. The main difference is that while the wage effect for second birth is negative and significant, the coefficient for the third birth is near zero and insignificant, indicating that the wage effect plays a marginal role in third births.

Looking more specifically at how wage may affect transition to second and third births, as reported in Table 1.4, we find the coefficients to be of similar intensity when controlling for wage and interaction with duration, region effects and institutional effects (see age at the first and second birth in Table 1.3 and Table 1.4). Wage has a negative impact on the probability of second birth one to two years after the first birth (column (1) of Table 1.4). This is either because households prefer to spend additional resources on the first child or because it is the moment women tend re-enter the labour market. Even if her wage increases, she has little incentive to use it for rearing another child. The situation is different after three to four years. In this case an increase in wage has a positive effect on second birth. A similar pattern is followed by women at risk of third birth (column (3) of Table 1.4, Wage$\times$ dur (3/4)=0.051). The wage effect is negative immediately after second birth and becomes positive (though not significant) three to four years after the second birth. This suggests that the timing for second and third births is positively induced by socioeconomic variables only three to four years after having experienced previous birth. Afterwards, the risk of giving a new birth decreases as we move away from the date of the preceding birth. This is perfectly in line with Heckman and Walker (1990b) point of view concerning 3rd birth in Sweden: when wages are higher there is a primary effect in reducing third birth and a secondary effect that forces pregnancy to be concentrated in a shorter span of life cycle.

Table 1.4 shows that women living in the Center and North are less likely to experience a second/third birth compared to the South. But Table 1.4 also reveals a clear institutional/cultural effect underlined by the interaction of wage and region. An additional unit of wage for women in the North increases her risk to have a second/third birth. One might argue that this reflects confidence in the child-care system among working mothers living in the North and Center (see del Boca, Locatelli and Vuri 2005 for a complete explanation). In Southern Italy, crêches are not widespread and working mothers tend to prefer informal

Table 1.4: Wage and Region effect for the transition to Second and Third birth.

Variable	Coefficient	(Standard Error)	Coefficient	(Standard Error)
	(1)	(2)	(3)	(4)
	Transition to second birth		Transition to third birth	
Age at first birth	-0.031**	(0.003)		
Age at second birth			-0.116**	(0.007)
Wage and duration:				
Wage ×dur (1/2)	-0.046†	(0.024)	-0.015	(0.064)
Wage ×dur (3/4)	0.069**	(0.017)	0.051	(0.044)
Wage ×dur (5/6)	-0.064**	(0.018)	-0.028	(0.052)
Wage ×dur (7/9)	-0.150**	(0.024)	-0.097*	(0.046)
Wage ×dur (10/13)	-0.168**	(0.044)		
Wage ×dur (10/14)			-0.328**	(0.063)
Wage ×dur (14/25)	0.130	(0.089)		
Wage ×dur (15/24)			-0.273	(0.220)
Region effects:				
Center	-0.432**	(0.031)	-0.304**	(0.081)
North	-0.489**	(0.025)	-0.248**	(0.068)
Wage and Region:				
Wage × Center	0.008	(0.031)	0.110	(0.086)
Wage × North	0.225**	(0.021)	0.274**	(0.058)
Constant	-2.842	(0.083)	-1.552**	(0.212)
No. of Observations	147,539		98,418	
No. of Women	25,227		14,483	
No. of Spells	14,334		2,506	
Pseudo R^2	0.091		0.049	
Log-L	-42686.643		-11089.553	

Notes: Discrete-time logit hazard regressions for the transition to second and third birth. Baseline not reported. Reference category for Region is South. Dur (1/2)= See section 1.3.3 and equation (1.3.5) in particular. The number of women at risk of third birth differs from the number of spells reported for the transition to second birth because women having twins as second births are not in the risk set of third birth. Bootstrapped standard errors in parentheses. Standard errors are computed using 300 bootstrap samples.

$^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$.

child care, i.e. baby-sitters and grandmothers living nearby. When any of these two conditions is not satisfied, having another child means additional costs in terms of searching for childminders. The decision to have a second or third birth, even if in presence of higher wage, is therefore not straightforward (see also Diprete, Morgan, Engelhardt and Pacalova 2003). Different is the situation in the North, where the large presence of public and private services offers a more reliable child-care system that also facilitates child-care even when children are very young (del Boca, Locatelli and Vuri 2005).

1.4.3 Turco's law (1999)

As we point out in the theoretical background, Italy has been one of the first countries to reach the so-called lowest-low fertility during the early 1990 (Kohler, Billari and Ortega 2002). Many researchers have focussed on the existence of an unmet need for family-friendly policies as one of the reasons behind lowest low fertility (see for example Demeny 2003). However, there is little scientific evidence concerning the impact of policies on fertility in a lowest-low setting. Moreover, there seems to be a general skepticism in the literature of whether public policies may have an impact on choices concerning fertility.

Table 1.5 considers only women being at risk of third birth, excluding 2003 because the last wave of the ISTAT-LFS took place in October 2003 and there are no data available for the births occurred in November and December 2003.

Generally women living in the South are more at risk of third birth when compared to the Northern ones. After the introduction of the law, the general period effect is negative confirming the lowest-low fertility for Italy. However, if we restrict the attention to the

Table 1.5: **Assessing the impact of Turco's Law (1999).**

Variable	**Coefficient**	**(Standard Error)**
Region (Ref. North):		
South	0.105^{*}	(0.053)
Center	-0.183^{*}	(0.071)
General trend (Ref. trend before the law):		
After 1999	-0.207^{**}	(0.047)
Poor after the law (Ref. poor women before 1999):		
Poor after 1999	$0.202^{\dagger}$	(0.120)
(*Not poor: Wage>15th perc. distribution*)		
Poor	0.356^{**}	(0.070)
Constant	-4.836^{**}	(0.080)
No. of Observations	86,572	
No. of Women	13,971	
No. of Spells	2,385	
Pseudo R^2	0.037	
Log-L	-10508.879	

Notes: Discrete-time logit hazard regression for the transition to Third Birth. Baseline not reported. Estimated equation (1.3.12). Poor women defined as women with wages less than the 15th percentile of the hourly wage distribution. 2003 excluded from the estimated equation. Bootstrapped standard errors in parentheses. Standard errors are computed using 300 bootstrap samples.

$^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$.

sub-group of poor women, it is clear that after 1999 they are at a higher risk of having a third birth when compared to the sub-group of poor women before the introduction of the law. Moreover, increasing the wage of a poor woman of one unit increases the probability of having a third child, and the effect is statistically significant at the 10% level (0.202). Cross tabulating the constructed poverty measure and the Region, we observe that 90% of the poorest women are living in the South. This confirms Official Statistics (Lelleri and Marzano 2002) showing that a significant share of households living in the South (see Table 1.6) received this transfer.

Table 1.6: Percentage of families receiving Turco's Law

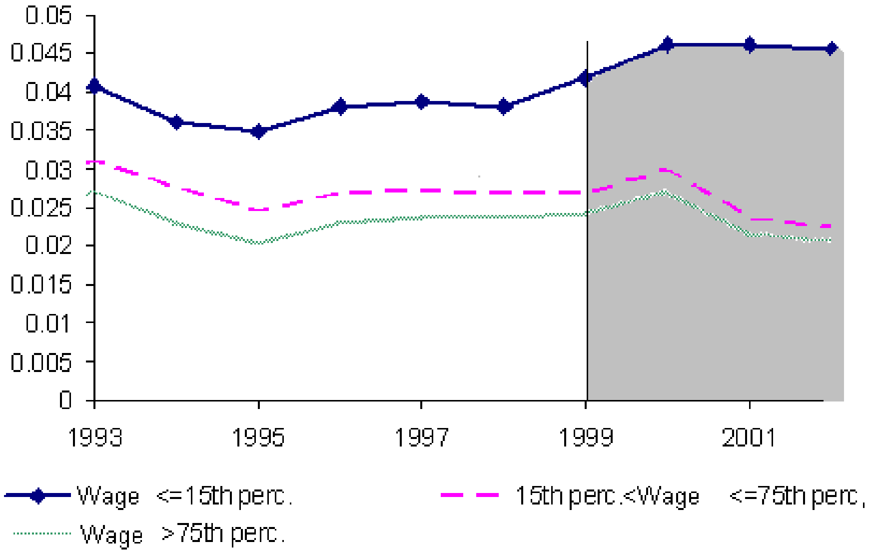
	1999	**2000**
Percentage of families with three or more children under 18:		
North	14.8	16.2
Center	21.7	23.7
South	58.8	64.2
Percentage of women receiving the second measure over total live births:		
North	11.9	12.3
Center	22.8	24.1
South	51.3	56.5

Source: Lelleri and Marzano, 2002

The impact of the law is, however, very much dependent on the specification of the poverty threshold. If we choose, for example, the 25th percentile of the distribution as threshold, the law does not seem to have much effect. The law could, reasonably, affect the fertility decision of very poor women because, only for this particular sub-group, the money transfer could produce a non-negligible proportion of the cost of having an additional child. But there can also be other factors (we cannot control for) influencing the positive coefficient for a poor woman after 1999, not related to the law. Indeed, Figure 1.7 shows that after 1999 fertility of poor women increased, but we cannot be sure whether this is due to the law or not.

In many respect the results are consistent with Gauthier (2001): "Overall, thus, the multivariate studies provide mixed conclusions as to the effect of policies on demographic and economic behavior, once other factors such as education, wage etc. are controlled for. The effect, if any, tends to be small. Methodological issues may be at the basis of these inconclusive findings...", (see also Gauthier 2004). We can argue that the Law had an impact,

Figure 1.7: Mean Predicted Hazard (1993-2002).



Notes: Fertility from 1993-2002. Mean predicted hazard per year of a discrete-time logit event history model (estimated from equation (1.3.12)) where wage has been stratified in low (less then 15th percentile of the distribution), medium (15th percentile< Wage<= 75th percentile), high (>=75th percentile). In grey the area where Turco's law could influence fertility.

but we are not able to estimate precisely the magnitude of its influence on increasing the rate of third births in Italy. The main difficulty lies with availability of data.

In order to asses the causal effect we need to use Regression Discontinuity techniques (or similar) that identifies the impact of the policy measure with the shift in the regression line before and after the law is introduced. The main problem is that our instrument (the wage) that discriminates between treated (who received the money) and not treated is estimated (see Appendix A) and we are not able to clearly identify the group of compliers (see Hahn, Todd and van der Klaauw 2002 for an explanation). This is because we need to identify the impact of the law on wage and how wage of poor households changed after receiving the benefit from the State.

1.5 Summary and Concluding Remarks

The aim of this chapter is to find an empirical connection between the striking increase in women's labour force participation and the delay in motherhood and the transition to second and third birth in Italy, though the impact of potential outcome.

Using two different data sets (from the ISTAT -LFS, 2003 and the BOI-SHIW, 2002), we estimate the effect of predicted wages on the postponement of motherhood. First, we estimate a wage equation including detailed controls for women's educational attainment using a Tobit model. Second, we use the predicted wage as a regressor in discrete time hazard models for the transition to first, second and third birth. We find that the wage effect is negatively correlated with having children. The magnitude, however, varies according to the birth order. Wage has a strong negative effect in the timing of first birth. Consistent with opportunity cost theories women with higher wage tend to delay motherhood. Our estimates suggests a non-proportional hazard and a strong, though incomplete, recuperation effect. Women with higher wage start having children later, but recuperate after some time, while poor women start earlier. Furthermore, there is no evidence that institutional or cultural effects are responsible for the postponement of maternity.

The pattern is different for second and third births. The wage effect is always negative, but it has smaller intensity compared to the first birth. The coefficient for third birth is close to zero, confirming that wage plays a relatively small role. Nevertheless, we find evidence of institutional differences in the decision of having more than one child. The empirical evidence for birth order two and three is in line with Ermisch (1989) indicating that the effect of women's wage depends on the availability of external child care. More specifically, an additional unit of wage for a Northern woman increase the risk of experiencing a second and third birth because they are more confident in the availability of a good child-care provision.

We conclude the empirical analysis by focussing on a public policy implemented in Italy in 1999 (Turco's law) which supported (through monetary transfers) households with children. Our estimates give some support for the law having an effect. But it is not clearly identifiable and quantifiable given our methodological approach, which in turn was driven by the data available. Despite this limitation our results provide interesting evidence of the key role wage plays in fertility decisions.

Chapter 2

Socio-economic Differences in Postponement and Recuperation of fertility in Italy: Results from a Multi-Spell Random Effect Model.

2.1 Introduction

It is a widely held view that postponement in fertility leads to a general decline in the total fertility rate (TFR). In other words, the longer a woman delays the onset of her child bearing career the lower is her completed fertility. This has been documented in various studies and the pattern has been accepted as an empirical regularity across cohorts and countries (see for instance Billari and Kohler 2002; Bumpass and Mburugu 1977; Bumpass, Rindfuss and Janosik 1978; Marini and Hodson 1981). This regularity has sparked considerable research into the determinants behind the timing of first births, arguing that a greater understanding of the driving forces behind the onset of the child bearing career, also informs us about the determinants of completed fertility. One of the important findings from this literature is that women with higher education or higher earnings potential tend to postpone first birth (Blossfeld and De Rose 1992; Marini 1984). However, the fact that women with higher education tend to delay first birth does not necessarily imply that they have lower completed fertility (Skirbekk, Kohler and Prskawetz 2004), or at least it is difficult to predict its exact extent. It is not unreasonable to think that the very same women who delayed the onset of childbearing, will also have an incentive to accelerate the second and third birth events in order to achieve their desired fertility level. This is commonly referred to as the recuperation effect on fertility.

Unfortunately the identification of postponement and recuperation effects on total fertility is not trivial. Whereas both will depend on observed characteristics, they will also depend on unobserved characteristics. Thus considerable caution is needed if one wants to establish the determinants and extent of the postponement and recuperation effects respectively. In light of these difficulties, the aim of this chapter is twofold. First we are interested in estimating the impact of women's social and economic characteristics on the decision to postpone fertility. Second, given differences in postponement, an interest lies in estimating the extent to which women with different characteristics are able to recuperate fertility within the first, second and third parities. The analysis enables us to assess to what extent socio-economic status might be driving the underlying low fertility in Italy. Our analysis is based on data from Italy, which was one of the first countries to experience lowest low fertility levels, i.e. TFR less than 1.3 (Billari and Kohler 2004). In 1995 Italian TFR fell to an all time low of 1.18, a development that has been followed by a general increase in the mean age at first birth. In order to estimate postponement and recuperation effects we specify a multi-spell random effect hazard model of the first, second and the third births. The model allows for unobserved heterogeneity, and we test to what extent the assumption about its functional form has any impact on the overall parameter estimates. Socio-economic status of women is inferred from the Bank of Italy Survey of Households' Income and Wealth (BOI-SHIW, 2002, from now on), whereas the model of fertility is based on the Italian Institute of Statistics Labor Force Survey of 2003 (ISTAT-LFS, from now on).

The chapter is organized as follows. In Section 2.2 we review the literature on the postponement effect focussing on the Italian context. Sections 2.3 and 2.4 describe methodological issues, including the surveys and the statistical model. In section 2.5 we discuss the empirical evidence concerning determinants of first, second and third births. Section 2.6 concludes.

2.2 Postponement and recuperation of fertility in Italy

Over the last two decades, the mean age at first birth increased substantially in many countries, a rise that has been linked to a significant decline in the observed total fertility rate. In Italy, mean age at the first birth (MAFB) was 25 years in 1980, 26.9 in 1990, 28.0 in 1995, 28.4 in 1996 and 28.7 in 1997 (Council of Europe 2004). During the same period the Total Fertility Rate (TFR) declined from 2.2 in 1980 to 1.18 in 1995 (see Table 2.1) which is below the threshold of 1.3 children per woman that Kohler, Billari and Ortega (2002) define as 'lowest-low fertility'. It seems therefore clear that the increase in mean age at first birth is associated with a substantial decline in the observed total fertility rate.

There are many explanations for why such a postponement has taken place. A prominent

Table 2.1: Education, Total Fertility Rate and Women Participation in Labor Market in Italy (1993-2000).

Variable	Period 1993	1994	1995	1996	1997	1998	1999	2000
Education	58.1	59.9	62.7	64.8	66.7	70	70.4	73.8
Employment Rate	35.8	35.4	35.4	36	36.4	37.3	38.3	39.6
Total Fertility Rate (TFR)	1.25	1.21	1.18	1.2	1.22	1.19	1.22	1.24

Notes : Fertility rates: 1998 and 1999 are provisional values and 2000 is an estimated value. Education is the percentage of the female population aged 20 to 24 having completed at least upper secondary education. The female employment rate is calculated by dividing the number of women aged 15 to 64 in employment by the total female population of the same age group.

Source: EUROSTAT. http://epp.eurostat.ec.europa.eu.

theory is that increased returns to female education and labour market participation makes childbearing more costly (Happel, Hill and Low 1984; Hotz and Miller 1988). Moreover, as women increase their participation in higher education, a delay of the onset of childbearing seems inevitable (Brewster and Rindfuss 2000; Cigno 1994). Gustafsson (2001) suggests that the delay is a likely consequence of the increasing presence of women in the labour force and of their level of human capital accumulation. This is particularly true of Italy. As Table 2.1 shows, the female employment rate increased during the 1990s, starting from a value of 35.4% in 1994 and reached 39.6% in 2000. During the same period, the percentage of women (aged from 20 to 24) having completed at least secondary school increased sharply from 58.1% in 1993 to a 73.8% in 2000. A related issue concerns increased difficulties in re-conciliating labour force participation and childbearing. Certainly, as women increase their commitment to work and educational attainment, child bearing becomes more strenuous if there is no follow-up in terms of availability of external child care. A further factor concerns lack of available and affordable housing. An interesting feature of Italian family dynamics is that young individuals are increasingly late in leaving the parental home, possibly due to lack of financial resources or affordable housing as Duce Tello (1995) suggests. The process of setting up one's dwelling can take several years after entering in the labor market. Additionally, Aassve, Billari, Mazzuco and Ongaro (2002) argue that the inherent uncertainty of the labor market perpetuates the pattern of staying at home with parents until a relatively high age, again having an effect on timing of marriage and the onset of child bearing.

It has been demonstrated that the youth labour market in Mediterranean Countries are considerably worse than many other European countries (Kohler, Billari and Ortega 2002). Low wage and a high prevalence of fixed term job contracts dominates, discouraging many young individuals from entering the labour market (see de la Rica and Iza 2005, for the Spanish case). Instead, entering higher education has become more popular and attractive. Certainly higher education has become an important way in which individuals can increase their chance of finding a stable job with a sufficient wage.

The empirical evidence on postponement and recuperation effects in Mediterranean countries based on micro level data are still limited. Using a simple regression framework on individual data, Kohler, Billari and Ortega (2002) find that in Italy the postponement effect is relatively high and it implies a relative reduction of completed fertility between 2.9% and 5.1% for each one-year delay in the onset of motherhood. For cohorts 1952-1958 they estimate a postponement effect equal to 2.9%, a level which is substantially above the levels found for Denmark and Sweden, two countries where there is evidence of strong recuperation. Applying an Heckman type sample selection model (Heckman 1979), Billari and Borgoni (2004) showed that the postponement effect is stronger in Italy than all other 'Southern European' lowest-low fertility countries. Overall they suggest that heavy postponement of the first birth and relatively low progression to higher level births is a common feature of these countries. Kohler, Billari and Ortega (2002) suggest that the lowest-low fertility in Italy is associated with a low progression probability to parities higher than the first, and that any recuperation effects are weak, and certainly incomplete.

One issue that complicates the analysis of postponement and recuperation is that it might be driven by unobserved factors. Marini and Hodson (1981) and Heckman, Hotz and Walker (1985) underline that only a part of the observed negative association between postponement and completed fertility is due to a causal mechanism; another part of the effect may be spurious. There might for instance be systematic differences in individuals' preferences for family size. If such preferences are not controlled for then estimates will be inconsistent and biased (Kohler, Skytthe and Christensen 2001). A strong preference for larger family size will naturally accelerate the first birth (Trussell and Menken 1978). But if such preferences are unobserved, then the impact of age at first birth on total fertility becomes overestimated. An additional source of unobserved heterogeneity may be differences in fecundity. Couples who have their first birth earlier may do so simply as a result of higher fecundity, whereas women with lower fecundity may conceive later. Importantly, these sources of heterogeneity might carry an effect onto subsequent births, leading to incorrect estimates of completed fertility. Yet another source of unobserved heterogeneity might be differences in ability which affects the incentives to invest in education and labour market skills. This in turn may affect age-related costs of postponing first birth (Kohler, Behrman and Watkins 2000). As will be clear from the next section, our data lack information about marriage and therefore information about husbands, which is another source of unobserved heterogeneity.

To overcome the estimation problems caused by unobserved factors, Kohler, Skytthe and

Christensen (2001) use Danish data on female monozygotic twins to estimate the postponement effect in a linear model for the logarithm of the total number of children. Nevertheless, there are two problems with this approach. First, a technical problem: the analysis on the logarithm of the total number of children is possible only when women are past or very close to the end of childbearing ages. Second, data on twins exist only in very few countries, and the postponement effect might depend on societal factors that vary across nations (Billari and Borgoni 2004; Kohler, Billari and Ortega 2002).

An alternative way to overcome these issues is to specify a parametric hazard regression model that explicitly controls for the influence of observed and unobserved heterogeneity. It is well known that when analyzing the timing of life-course events, failure to control for sample heterogeneity may produce severe biases in structural estimates of duration models (see for instance Heckman and Singer 1984; Lancaster 1990; Vaupel, Manton and Stallard 1979). The choice of the heterogeneity distribution might also matter. Heckman and Singer (1984) argue for instance that an incorrect assumption on the distribution of the unobserved heterogeneity may cause serious parameter bias, and suggest as a result the use of a flexible non-parametric distribution of unobserved heterogeneity. In this chapter, we specify a discrete time regression model of first, second and third births, estimated in a joint framework. The model is estimated with different assumptions concerning the heterogeneity component, and we assess to what extent different assumptions impacts the parameter estimates of interest.

2.3 Data

The socio-economic status is measured by women's predicted earnings. Unfortunately, none of the Italian data sets currently available contains information on both fertility histories and measures of their opportunity cost (i.e. their wage and income). Wage information from the the Bank of Italy Survey of Households' Income and Wealth, 2002 (BOI-SHIW) is merged with fertility information from the the Italian Institute of Statistics Labor Force Survey, 2003 (ISTAT-LSF). The benefit of the BOI-SHIW is that it contains detailed information on earnings and income. A serious disadvantage is that the sample size is small, a drawback that becomes particularly critical for the estimation of third birth events. Consequently we use the BOI-SHIW to estimate wage and earnings equations, where predicted wages are matched with individuals in the ISTAT-LFS data set. In brief, a Tobit augmented log-wage equation is estimated as follows:

$$\hat{\omega}_{it} = \hat{\beta}_0 + \hat{\beta}_1 \text{age}_{it} + \hat{\beta}_2 (\text{age})^2_{it} + \hat{\beta}_3 \sum_{j=1}^{8} \text{education}_{itj} + \hat{\beta}_4 \sum_{j=1}^{20} \text{region}_{itj} \tag{2.3.1}$$

where $\hat{\omega}_{it}$ denotes the log hourly predicted wage for the $i-th$ woman at time t, $t = 1983, ..., 2003$, age_{it} is the age of the $i-th$ woman at time t, $(\text{education})_{itj}$, for $j = 1, ..., 8$, are eight dummy variables for different levels of education attained from the $i-th$ woman at time t and $(\text{region})_{itj}$, for $j = 1, ..., 20$, are twenty dummy variables, one for each of the Italian regions (for a detailed discussion see appendix A). In line with the economic literature, wage increases with age but at decreasing rate (Mincer 1974). Education is an important driver of wage levels; the higher the education level, the higher the wage (Schultz 1985). The hourly wage does not exhibit very large variability across different regions of Italy, but there is a clear North/South divide. This difference is caused by prevalence of public jobs and the percentage of women not working in the South (see appendix A, Table A.1).

Our multi-spell random effect model (estimated from ISTAT-LFS) uses wages predicted from the BOI-SHIW. In order to take into account the fact we have generated regressors (Arellano and Meghir 1992; Pagan 1984), the standard errors are corrected bootstrapping the 20% of the original sample using 100 replications with observations clustered at the woman level.

The woman is taken as the unit of analysis and linked with her co-resident children at the moment of the interview in the ISTAT-LFS. In order to ensure that the recorded children are the only ones of the mother, we limit the analysis to only include women who are 40 or less in 2003 and we reconstruct retrospectively their fertility histories going back to 1983. We also drop all households where we were unable to link children with mothers (i.e. male head of the household with no wife and all single men). In total, the fraction of women that can be matched with their co-residing children is 95%. Note that we are unable to use information of husbands since we only know the women's marital status in 2003. Thus, marital histories cannot be reconstructed.

Following this approach we end up with a sample of 34,439 women, for which we have information about their childbearing career. Every woman is right censored in 2003 or if she experience all the three births at time, say t, we observe her until that time t. If she gives a birth of parity j in time T she is at risk for a new birth since $T + 1$ and until she experience a new birth $j + 1$ or, if she does not, until 2003. All women are followed from age 15, which we assume is the start of her childbearing career. In Table 2.2 we report basic descriptions of the sample across different parities.

Table 2.2: Frequency Table of Women at Risk of First, Second and Third Birth.

Parity	Women at Risk of having a child of parity j	Women having a Child of parity j
First Birth $j = 1$	34,224	25,281
		(73.86%)
Second Birth $j = 2$	25,275	14,334
		(56.71%)
Third Birth $j = 3$	14,508	2,506
		(17.27%)

Notes: Women aged from 15 to 40 in 2003 included in the sample. The number of women at risk of a second birth differs from the number of those having a child of parity one because women having twins as first birth are directly at risk of a third birth. For the same reason, women having twins as second birth are not in the risk set of third birth.

Source: Own Calculation from ISTAT-LFS, 2003.

2.4 The Model

We present two versions of the model. The first assumes that unobserved heterogeneity can be captured by a normally distributed error term. The second assumes a non-parametric specification for the unobserved women-specific characteristics. The second version is similar to Heckman and Singer (1984), but differently from that approach we specify the baseline hazard to follow a piecewise linear spline specification.

We model the decision to have the first, second and third birth as a multiple spell duration model:

$$H^*_{ijt} = \gamma_j \left(T_j(t)W_{ijt}\right) + \beta_j^{'} X_{ijt} + \mu_i + \epsilon_{ijt} \qquad (2.4.2)$$

where H^*_{ijt} indicates the propensity of the $i-th$ woman to give $j-th$ birth ($j = 1, 2, 3$ for first, second and third birth) at time t, $T_j(t)$ is a spline for parity j, W_{ijt} indicates the wage while X_{ijt} captures observed characteristics other than wage for the $i-th$ woman at time t for parity j, μ_i represents the random effect specific for each woman and ϵ_{ijt} is a residual error term distributed as a logistic with zero mean and variance $\pi^2/3$. The most aggregated level is the woman, and therefore the unit of analysis. Since each woman might have more than one child, we can consider parity $j = 1$ if the woman is at the risk of having the first child, $j = 2$ if first birth has taken place and she is at risk of having the second and $j = 3$ if the woman experienced a second birth, so that she is at risk of having a third child. Time periods are the most disaggregated level $t = 1, ..., T$, and all time varying variables are given at this level. From equation (2.4.2) it is clear that if $H^*_{ijt} < 0$, then the woman does not give the $j - th$ birth and $H_{ijt} = 0$. If $H^*_{ijt} \geq 0$ she experiences the $j - th$ birth as $H_{ijt} = 1$.

Duration dependence for first, second and third births are modeled separately. Importantly, women's wage levels (W_{ijt}) are interacted with the splines ($T_j(t)$). This means that the resulting parameter estimates gives us direct measures of the extent to which women with different wage levels may postpone the first birth, and possibly recuperate through second and third births. Within each parity order we stratify the wage variable into low, medium, high and very high wage. The stratification was chosen according to the 25th, 50th and 75th percentile of the wage distribution. Other observed characteristics at the woman and child level (across time) are captured by X_{ijt}, and this includes region, age at the first birth, age at the second birth. The subscript of β_j signifies that women living in the same region, having a first birth at the same age may behave differently depending on the parity considered.

Unmeasured characteristics (μ_i) are assumed to be woman-specific and constant across time and parity orders. An important question in this analysis concerns the amount of heterogeneity captured by the error term μ_i. Given our data source, there are certainly many reasons why we could expect it to be important. For instance, the lack of data on marriage histories prevents us from reconstructing retrospective marriage histories, which in effect leaves out any relevant information about husbands' role in fertility outcomes.

We study the sensitivity of the model to parametric and non parametric specification for the distribution of the unobserved heterogeneity. We start by assuming a normal specification for the random effect, $\mu_i \sim N(0, \sigma_\mu^2)$, and compare this model to one without random effect. We then extend this parametric setting with a non-parametric one specifying two and three mass points. There is no closed form solution of the likelihood function for hazard models with random effects. The residual or the random effect is therefore integrated out using a numerical integration algorithm based on Gauss Hermite Quadrature (Abramowitz and Stegun 1972). This algorithm selects a number of support points and weights such that the weighed points approximate a normal distribution:

$$\int_{-\infty}^{\infty} \phi(\epsilon) L(\epsilon) dx \approx \sum_{l=1}^{k} w_l x_l \tag{2.4.3}$$

where w_l and x_l are Gauss-Hermite weights and support point, respectively. The higher the number of support points the more accurate the approximation, but the slower the computation [1]. The Non Parametric Maximum Likelihood (NPML) is highly sensitive to starting values and estimation is difficult when the location of the mass points were specified arbitrary. Estimation is considerably easier when the unobserved heterogeneity is specified by the normal distribution. It is therefore convenient to estimate the model with normally distributed errors, and take these estimates as starting values for the non-parametric version. In the case of the non-parametric specification we can think to μ_i in terms of a "mover-stayer" structure. The "movers" reflect individuals who, for some unobserved reason, make the transition faster than the "stayers". Because unobserved heterogeneity may arise from

[1] In the case of the finite mixture Unobserved Heterogeneity we chose 10 integration points.

a number of sources, any specific interpretation is difficult, though in the next section we provide a tentative interpretation of the estimated masspoints.

Having estimated these models we offer an extension to model (2.4.2) in which the regressors are interacted with the unobserved heterogeneity term. For reasons will become clear in next section, the analysis is conducted when the unobserved heterogeneity term is specified parametrically (i.e. normal). In particular we estimate four models, one for each interaction of the normal random effect with a different wage level:

$$H^*_{ijt} = \gamma_j \left(T_j(t) W_{ijt}\right) + \beta_j^{'} X_{ijt} + \mu_i W_l + \epsilon_{ijt} \tag{2.4.4}$$

where W_l indicates wage levels, say $l = 1$ low, $l = 2$ medium, $l = 3$ high, $l = 4$ very high wage. With a variance decomposition we assess which socio-economic group captures the largest amount of the unobserved heterogeneity. If we denote $\mu_i W_l = u_{il}$ a normal random effect we get:

$$\mathbb{V}ar(u_{il}) = \mathbb{V}ar(\mathbb{E}(u_{il}|l)) + \mathbb{E}(\mathbb{V}ar(u_{il}|l)) \tag{2.4.5}$$

and because the random effect is assumed to be normally distributed with zero mean and variance to be estimated, the right hand side of (2.4.5) reduces to $\mathbb{E}(\mathbb{V}ar(u_{il}|l))$. More precisely:

$$\begin{aligned} \mathbb{E}(\mathbb{V}ar(u_{il}|l)) &= E_l(\sigma_l^2) \\ &= \sum_{s=1}^{4} \sigma_s^2 \, Pr(l = s) \end{aligned} \tag{2.4.6}$$

where $Pr(l = s)$ if $s = 1$ is the proportion of poor women in the sample. As a consequence, the part of variance captured by a low wage woman is going to be:

$$\frac{\sigma_1^2\, Pr(l=1)}{\sum_{s=1}^{4} \sigma_s^2\, Pr(l=s)} \tag{2.4.7}$$

Next section presents the main results concerning the models introduced in (2.4.2) and (2.4.4) and we offer an explanation of the main income and recuperation effect together with the role played by the age at the first birth.

2.5 Results

We start by discussing the parameter estimates for first, second and third births respectively. We then assess the estimates when heterogeneity is included compared to when it is not. The model without unobserved heterogeneity is presented in column (1) of Table 2.4, the estimates where unobserved heterogeneity is included and assumed normal is presented in column (2) in the same Table, whereas the models with non parametric unobserved heterogeneity are presented in columns (3) and (4). Finally we summarise the results when wages are interacted with the heterogeneity term.

2.5.1 First Birth

From age 15 to 18 women of different wage groups are quite similar in terms of their likelihood of entering motherhood, as illustrated in Figure 2.1 plotting the coefficients estimated in column (2) of Table 2.4. It is from the age 18 and onwards the rate of first birth starts to diverge, and the pattern is clear: those with low wages have a considerably higher rate of entering motherhood compared to those with high wages. The difference reflects a clear

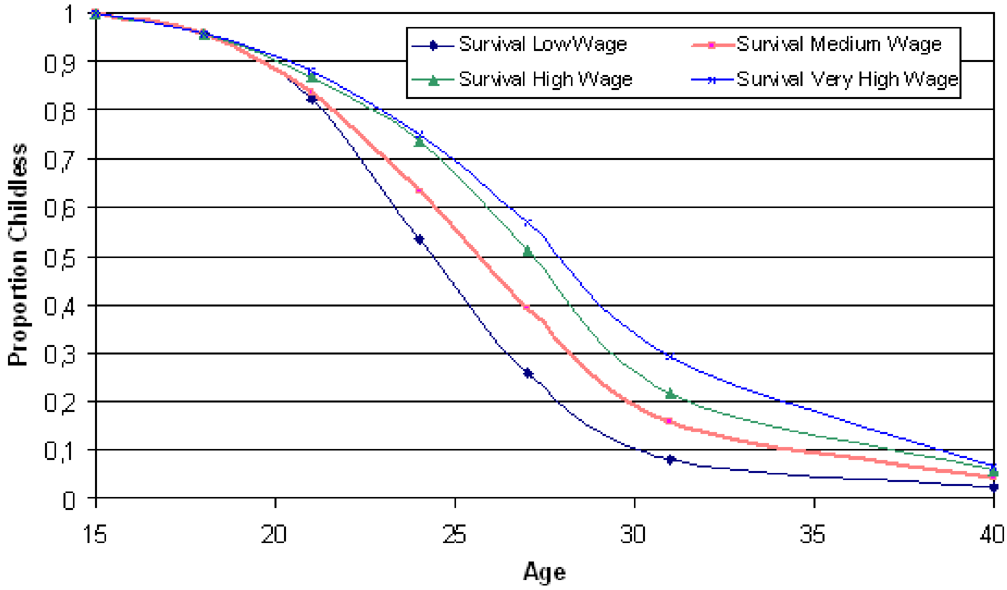
Figure 2.1: Estimated Spline for First Birth Across Different Wage Levels

Notes: Logit(hazard) estimated from Table 2.4, column (2), panel for First Birth. Wage has been stratified in Low (Wage less than 25th percentile of the Hourly Wage Distribution), Medium (25th percentile< Wage<= 50th percentile of the Hourly Wage Distribution), High (50th percentile< Wage<= 75th percentile of the Hourly Wage Distribution) and Very High (Wage>75th percentile of the Hourly Wage Distribution).

postponement effect, in that those staying on in higher education (i.e. those with high predicted wages) tend to delay the onset of childbearing. The differences between wage groups are significant until age 32 where the hazard rates cross. From this age onwards, those with low wage have the lowest rate of entering motherhood.

Postponement and recuperation effects can be better viewed through the predicted survival curves. These are plotted in Figure 2.2. The postponement effects are clear, as those with lower wages have a considerably steeper survival curve (i.e. they have the first birth earlier). However, the most remarkable feature of Figure 2.2 is the very strong recuperation effect. By age 40 the proportion of women with high wages having had the first birth is almost the same as for women with low wages. It is worth bearing in mind that these are the predicted survival curves based on our parametric model. Also note that in our sample we have omitted women aged 40 and over, thereby leaving out many who are recorded as childless, which in turn explains why the predicted survival curve is close to zero at age 40.

Figure 2.2: Estimated Spline-Survival Curves for First Birth Across Different Wage Levels



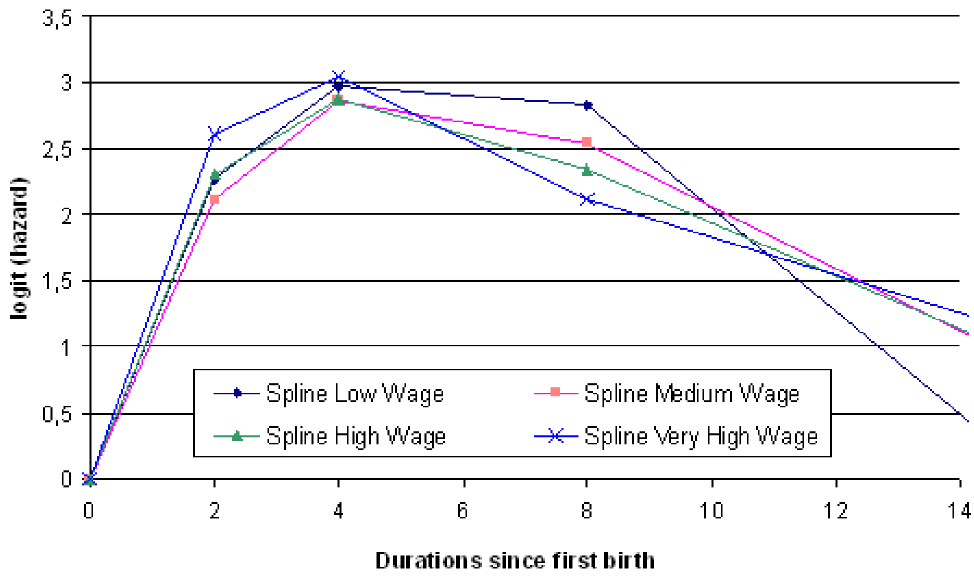
Notes: Survival Curves estimated from the coefficients reported in Table 2.4, column (2), for First Birth. Wage has been stratified in Low (Wage less than 25th percentile of the Hourly Wage Distribution), Medium (25th percentile< Wage<= 50th percentile of the Hourly Wage Distribution), High (50th percentile< Wage<= 75th percentile of the Hourly Wage Distribution) and Very High (Wage>75th percentile of the Hourly Wage Distribution). Women aged 40 and over are omitted from the sample: the predictions are made for women having children and ignoring censoring, i.e. women who remain childless. This explains why the predicted survival curves go so close to zero.

2.5.2 Second Birth

Figure 2.3 plots the logit(hazard) for second birth as estimated in column (2) of Table 2.4.

Here we see that women with very high predicted wages show a higher logit(hazard) from the beginning. However, the risk is only higher until 4 years after the first birth takes place, and after five years they have the lowest, whereas after 12 years women with the lowest wage have the lowest risk of second births. Thus, the effect of wage is mixed in that the

Figure 2.3: Estimated Spline for Second Birth Across Different Wage Levels



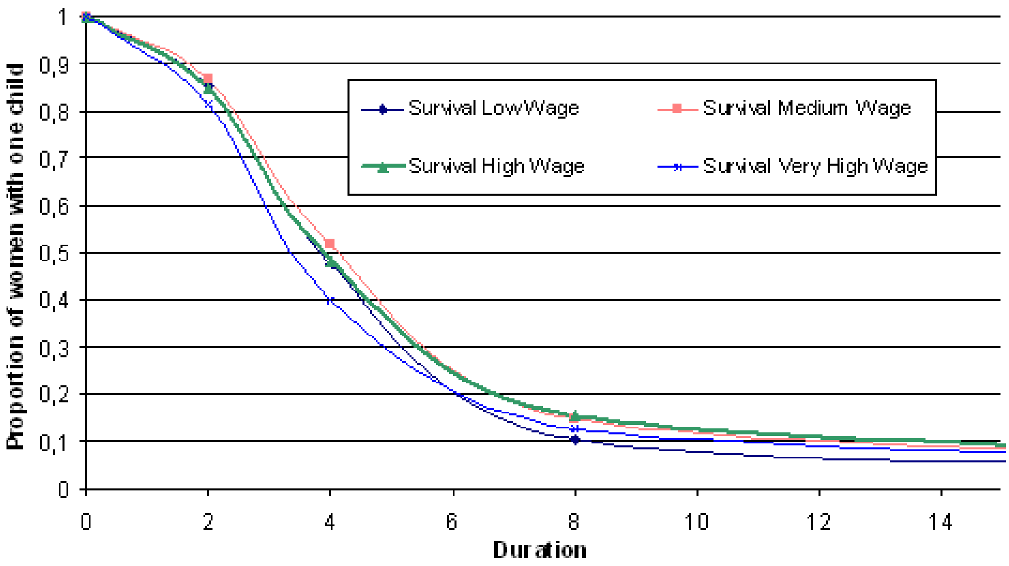
Notes: Logit(hazard) estimated from Table 2.4, column (2), panel for Second Birth. Wage has been stratified in Low (Wage less than 25th percentile of the Hourly Wage Distribution), Medium (25th percentile< Wage<= 50th percentile of the Hourly Wage Distribution), High (50th percentile< Wage<= 75th percentile of the Hourly Wage Distribution) and Very High (Wage>75th percentile of the Hourly Wage Distribution).

relative rank between in the groups (in terms of wage level) varies by duration. The fact that women with high wages have the highest hazard in the beginning is consistent with the argument of Heckman and Walker[2] (1990b) who suggest that women with high wages tend to concentrate their fertility history in a shorter period of their life. In this sense, we observe here a recuperation effect among women with the highest wages.

The pattern is clearly confirmed in Figure 2.4 where the predicted survival curves for women with high wages lies below the other groups. After two years the proportion of those with highest wages not having another child is 0.81, whereas this proportion falls sharply to 0.40 three to four years after they entered motherhood. In summary, our results show that

[2]They apply this argument to third birth in Sweden.

Figure 2.4: Estimated Spline-Survival Curves for Second Birth Across Different Wage Levels



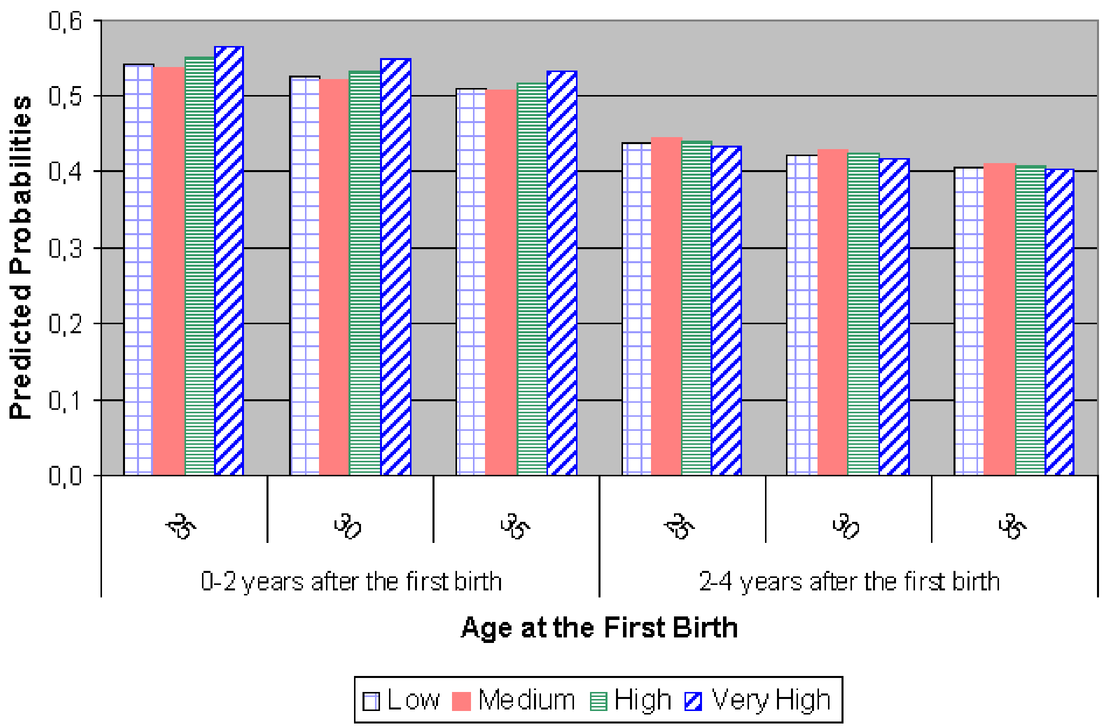
Notes: Survival Curves estimated from the coefficients reported in Table 2.4, column (2), for Second Birth. Wage has been stratified in Low (Wage less than 25th percentile of the Hourly Wage Distribution), Medium (25th percentile< Wage<= 50th percentile of the Hourly Wage Distribution), High (50th percentile< Wage<= 75th percentile of the Hourly Wage Distribution) and Very High (Wage>75th percentile of the Hourly Wage Distribution).

whereas those with high predicted wages tend to postpone the onset of childbearing, they also recuperate, not only through first birth, but also through the second. As was clear, the recuperation was not complete during parity one, but recuperation is close to complete once considering second birth. In other words, Figure 2.4 underlines that the recuperation is delayed and only completed through the second parity.

Of course, the explanation is quite clear. For women to reach a higher economic status they need to invest longer in education, and therefore postpone the first birth. Once education is completed, they gradually recuperate compared to the other women. In order to recuperate completely they also need to accelerate second birth compared to the others.

In this way they may also exploit the scale economies: having two children of similar age concentrates and possibly shortens the time needed away from work for the purpose of childrearing. Obviously, one can make the opposite argument: births occurring within shorter amount of time may result in an increase in physical and psychological costs in infant care (Bumpass and Mburugu 1978).

Figure 2.5: Probability of a Second Birth Across Different Wage Levels.



Notes: Predicted Probabilities from Table 2.4, column (2), panel for Second Birth. Probabilities estimated conditioning on age at the Firth Birth. Low, Medium, High and Very High refer to the wage stratification: Low (Wage less than 25th percentile of the Hourly Wage Distribution), Medium (25th percentile< Wage<= 50th percentile of the Hourly Wage Distribution), High (50th percentile< Wage<= 75th percentile of the Hourly Wage Distribution) and Very High (Wage>75th percentile of the Hourly Wage Distribution).

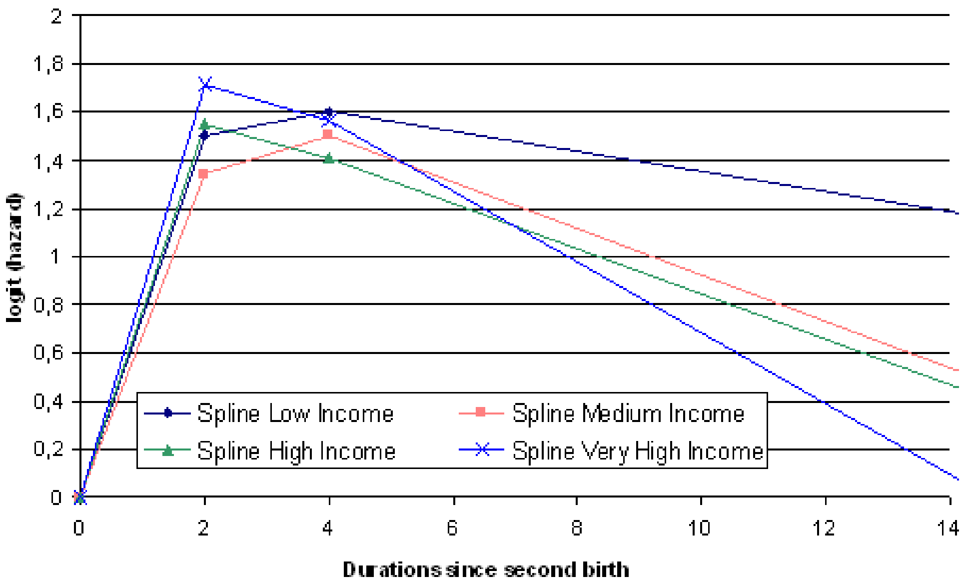
In Figure 2.5 we have plotted the predicted probability of having a second birth conditioning on the age of the first one with coefficients estimated as showed in column (2) of Table 2.4. Age at the first birth has a negative effect on the pace of subsequent fertility: this effect is particularly strong in the two years following the first birth and persists (with lower probability) even after four years the first birth happened. The experience of early motherhood tend to be associated with an increase in the probability of a second birth, so that early fertility is associated with a rapid and subsequent fertility. For a woman with low predicted wage the probability of having a second birth in the two years after she gave the first birth and conditioned on the fact she had the first one when she was 25, approaches 0.54. However if another low wage woman had the first child when she was 35, the probability of giving a second birth in the two following years is lower (around 0.50). The probability of giving a second birth two to four years after a high wage woman had the first one are 0.44, 0.42 and 0.40 if she gave the first birth at the age of 25, 30 and 35 respectively. In particular, for women having a second birth four years after they experienced the first one, there is a clear homogenous pattern of these probabilities across different wage levels. However a formal likelihood ratio test reject the null of no differences, suggesting there is an improvement in the model if we stratify the wage in four categories [3].

2.5.3 Third Birth

The splines plotted in Figure 2.6 (resulting from the estimated coefficients of Table 2.4, column (2)) suggest that for short durations, i.e. within the forth year since the second birth, a third birth is less dependent on wage level. The only exception is women with very high income and who conceived in the interval immediately following the second birth.

[3] More formally, we find that when collapsing women with low and medium wages into one category and comparing the model with the one reported in column (2) of Table 2.4, the difference in the deviance statistics (2×133.67) exceeds the 0.1% critical value of a χ^2 distribution with 13 degrees of freedom (34.52). A similar argument applies if we collapse women of high and medium wage: the deviance statistics assumes the value of 2×151.44 that again exceeds the mentioned percentile of a χ^2 distribution with 15 degrees of freedom. When considering into one category women with high and very high wages the deviance is equal to 2×86.76 and the rejection of the null applies again.

Figure 2.6: Estimated Spline for Third Birth Across Different Wage Levels



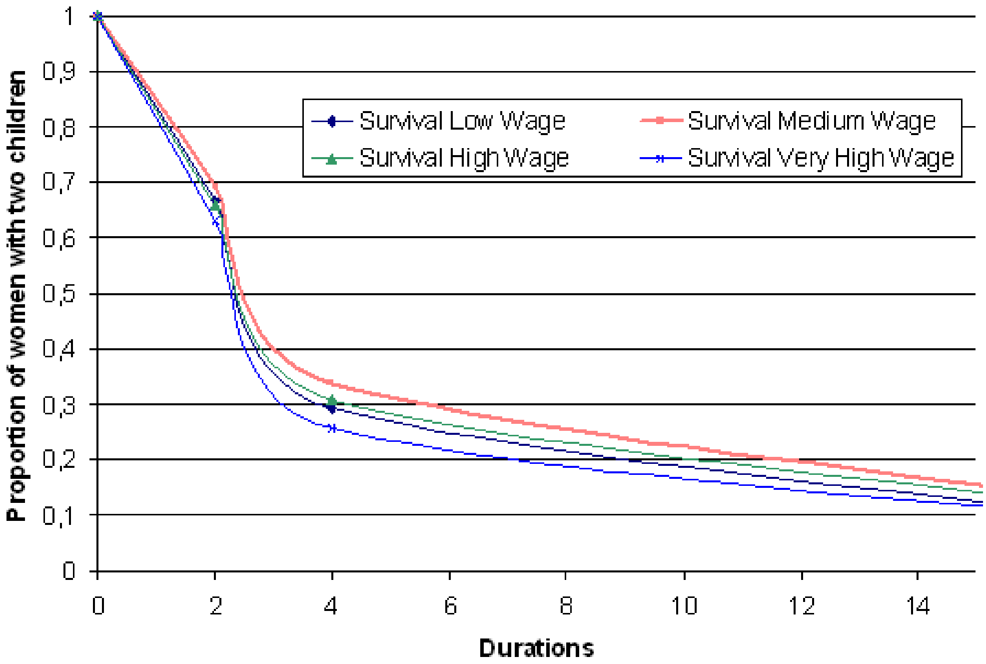
Notes: Logit(hazard) estimated from Table 2.4, column (2), panel for Third Birth. Wage has been stratified in Low (Wage less than 25th percentile of the Hourly Wage Distribution), Medium (25th percentile< Wage<= 50th percentile of the Hourly Wage Distribution), High (50th percentile< Wage<= 75th percentile of the Hourly Wage Distribution) and Very High (Wage>75th percentile of the Hourly Wage Distribution).

However, for very long durations the scenario changes and women with low wages remain at higher risk of experiencing a third birth given they have already had the second one.

The survival curves in Figure 2.7 show that there is little difference between the different wage groups, apart from an evident distinct pattern after four years the woman conceive the second birth. At that time there is a high proportion of women with medium wage remaining with two children, while the proportion of those with very high income and two children is smaller. It seems clear therefore that socio-economic factors do not play any important role in explaining why women choose to have third child.

We have so far considered the role of socio-economic factors, measured through their predicted wages. When considering other variables, we find that women living in the South

Figure 2.7: Estimated Spline-Survival Curves for Third Birth Across Different Wage Levels



Notes: Survival Curves estimated from the coefficients reported in Table 2.4, column (2), for Third Birth. Wage has been stratified in Low (Wage less than 25th percentile of the Hourly Wage Distribution), Medium (25th percentile< Wage<= 50th percentile of the Hourly Wage Distribution), High (50th percentile< Wage<= 75th percentile of the Hourly Wage Distribution) and Very High (Wage>75th percentile of the Hourly Wage Distribution).

have a higher rate of childbearing in all parities (see Table 2.4). The Center and North are very similar and the estimated coefficients interacting region and wages for the Center is close to zero. Among women living in the North, those with high wages have a higher risk of second and third births. This is in line with the argument of Ermisch (1989) who suggest that women with very high income are more able to afford external childcare, and thereby generating higher rates of childbearing. This argument applies mostly to second and third birth, while for first birth the introduction of the random effect captures parts of the effect.

2.5.4 Unobserved Heterogeneity

The main consequence from introducing a random effect on first birth is to increase the estimated parameters for women in the low and medium wage groups who make the transition after age 24 (see Table 2.4). For those entering motherhood after 27 we see an impact for those having high and very high wages. For the transition to second birth the effect of the random effect is to increase the estimated parameters four to eight years after the first child is born. This is the case for low, medium and high wage women. Women with very high income, instead, experience this effect after two to four years they enter into motherhood. This is a consequence of the fact that women with very high income tend to have the second birth rapidly after they had the first. The random effect also captures the effect of postponed marriage, since after marriage they tend to concentrate their childbearing in a shorter time span. The main consequence of introducing the random effect for third births is in the intercept of the process as column (2) of Table 2.4 shows.

The random effect may explain the relative inconvenience for a woman to take time away from work when the previous birth happened in the past. For women with very high income, instead, unobserved factors could be linked to the higher likelihood of having another birth as soon as you had the first one. Without particularly focussing on socio-economic features, Marini and Hodson (1981) show that the spacing of the first birth have a causal effect on the spacing of the second.

As pointed out at the beginning of the section, the functional form of the unobserved heterogeneity does not impact on the estimated coefficients even when a three mass point specification is assumed. Despite this, the model is sensitive to the introduction of the random effect. The standard deviation of the unobserved heterogeneity is 0.53 when it is modeled as a normal distribution (column (2) of Table 2.4). When we estimate the NPML with two mass points we find that the two groups are almost identical: 0.54% and 0.46% of the women (see column (3) of Table 2.4). The standard errors, however, do not differ much from the normal case (0.58 for two mass points). The algorithm also converges with three mass points (column (4) of Table 2.4), finding one large group (64%), one medium group (33%) and one very small group (3%). The variance explained by a three mass point non parametric finite mixture is close to the one with two mass points.

The interpretation of the mass points, of course, is difficult and any explanation is merely tentative. In applications of fertility behavior they are often taken to reflect differences in family orientation. Thus women with a strong family orientation (more likely to marry and have children) would be associated with a positive masspoint ("movers"). The group with a negative masspoint (the "stayers") can be thought of women with a higher career orientation, as it is natural to assume that have a lower rate of marriage and childbearing as a result. Of course, it is also possible that the masspoints somehow reflect our omission of the husband's wages as covariates in the fertility process.

2.5.5 Variance Decomposition

We now extend this model as suggested in equation (2.4.4). In particular we would like to asses which wage level captures the largest part of the variance of the random effect. To this extent we estimate four separated models which differ by the interaction of the random effect with a particular wage level (i.e. socio-economic group) (see Table 2.3). More precisely the

Table 2.3: Estimated Standard Errors of the Random Effect When Interacted With Different Wages Levels.

Variables	Normal Random Effect (1)		$Pr(l = s)$: Frequency (2)
Without Interaction	0.534***		
Low Wage × Random Effect	0.346***	$s = 1$	0.25
Medium Wage× Random Effect	0.185**	$s = 2$	0.25
High Wage× Random Effect	0.382***	$s = 3$	0.25
Very High Wage× Random Effect	0.222**	$s = 4$	0.25

Notes: Estimates of the other covariates included in model (2.4.4) are omitted but are not significantly different from the one reported in Table 2.4. $s = 1, ..., 4$ refer to four estimated models. Variables indicate the level of wage interacted with the random effect. Column (1) reports the estimates of the standard errors of the random effect for the related model. $Pr(l = s)$ for $s = 1, ..., 4$ is the proportion of women with low, medium, high and very high wage respectively.

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

amount of explained variance is 0.34, 0.09, 0.41 and 0.14 for the group of women with low, medium, high and very high income, respectively. This implies that women with relatively low and high wages are the typology of women for which it is more difficult to explain their attitude towards fertility decisions without considering other unobserved characteristics.

A tentative interpretation of this result can be the role of a childcare availabilities as suggested by Ermisch (1989). To this extent women with high wages are more likely to have children because they can afford external childcare. Women with low wages, instead, have more children because of their low opportunity cost and they are more family oriented (Becker 1965; Becker and Lewis 1973). In this scenario only those with medium and very high wages have less children. The former because they cannot afford external childcare, the latter because they are highly career oriented. As a consequence, the interacted wage random effect could capture the remaining effect of the socio-economic environment not entirely explained by wage.

2.6 Conclusions

The focus of this chapter has been to investigate socio-economic differences in delay and recuperation of childbearing in Italy, a country that is suffering from what is termed the lowest-low fertility. In order to do so we have combined information from the BOI-SHIW and the ISTAT-LFS. The former is used to collect information about women's socio-economic status, measured in terms of their predicted wages, whereas the latter has provided us with sample sizes ensuring safe estimation of birth parities. The approach is based on a hazard regression, whereby estimation of first, second and third births is done jointly. As a result we have been able to control for individual specific unobserved heterogeneity for which we have experimented with alternative assumed distribution functions.

Our main finding is that there are strong differences in the onset of childbearing or the timing of first birth. Women with high wages delay this transition considerably compared to women with low predicted wages. This is not unexpected as the strongest predictor behind earnings is educational attainment. Thus women with high earnings spend longer time in education, and possibly in the labour market, before they start their childbearing career. However, we also observe a strong recuperation effect, and by the age of 40 high earning women have caught up with low earning women almost completely. The recuperation of high earning women continues into the progression of second birth. Interestingly there are no strong socio-economic differences in terms of progression to third births. The result is highly interesting, since it suggest that any delay in child bearing due to higher predicted earnings, cannot explain the underlying low fertility levels in Italy. Other forces are at play.

Our analysis has also provided a detailed sensitivity analysis of the possible roles played by the assumptions imposed on the statistical model. We show that, though the introduction of unobserved heterogeneity is important, the assumption of its functional form is not. This is most likely a result of the fact that the baseline hazard is given a flexible functional form, and that predicted wages are interacted with the unobserved heterogeneity term.

Table 2.4: Estimated Coefficients for First, Second and Third Birth without Random Effect, with Normal Random Effect and a Non Parametric Specification of the Unobserved Heterogeneity with Two and Three Mass Points.

	Without RE	Normal RE	Two Mass Points	Three Mass Points
Variables	(1)	(2)	(3)	(4)
First Birth:				
Constant	-4.5359 ***	-4.6622 ***	-4.3481 ***	-4.7090 ***
	(0.1719)	(0.1771)	(0.3834)	(0.8859)
Spline1×Low Wage	0.2570 ***	0.2635 ***	0.2616 ***	0.2646 ***
	(0.0691)	(0.0697)	(0.0698)	(0.0656)
Spline2×Low Wage	0.5068 ***	0.5179 ***	0.5145 ***	0.5204 ***
	(0.0403)	(0.0402)	(0.0404)	(0.0370)
Spline3×Low Wage	0.2039 ***	0.2416 ***	0.2393 ***	0.2445 ***
	(0.0441)	(0.0472)	(0.0460)	(0.0426)
Spline4×Low Wage	0.0734 ***	0.1214 ***	0.1242 ***	0.1332 ***
	(0.0557)	(0.0589)	(0.0643)	(0.0574)
Spline5×Low Wage	-0.0491	-0.0087	-0.0120	0.0129
	(0.0758)	(0.0769)	(0.0895)	(0.0773)
Spline6×Low Wage	-0.2208 **	-0.1978 *	-0.2066 *	-0.1843
	(0.3844)	(0.3889)	(0.3910)	(0.3524)
Spline1×Medium Wage	0.3271 ***	0.3352 ***	0.3345 ***	0.3380 ***
	(0.0686)	(0.0693)	(0.0671)	(0.0667)
Spline2×Medium Wage	0.3007 ***	0.3145 ***	0.3112 ***	0.3173 ***
	(0.0592)	(0.0581)	(0.0589)	(0.0554)
Spline3×Medium Wage	0.1864 ***	0.2092 ***	0.2059 ***	0.2108 ***
	(0.0406)	(0.0416)	(0.0417)	(0.0373)
Spline4×Medium Wage	0.1267 ***	0.1605 ***	0.1614 ***	0.1643 ***
	(0.0468)	(0.0486)	(0.0507)	(0.0464)
Spline5×Medium Wage	0.0235	0.0679 ***	0.0692 ***	0.0853 ***
	(0.0554)	(0.0548)	(0.0633)	(0.0564)
Spline6×Medium Wage	-0.1973 ***	-0.1657 ***	-0.1751 ***	-0.1467 ***
	(0.0953)	(0.0884)	(0.1013)	(0.0970)
Spline1×High Wage	0.3409 ***	0.3509 ***	0.3506 ***	0.3573 ***
	(0.0880)	(0.0867)	(0.0875)	(0.0816)
Spline2×High Wage	0.0590 *	0.0781 **	0.0748 **	0.0794 **
	(0.0839)	(0.0807)	(0.0829)	(0.0773)
Spline3×High Wage	0.2849 ***	0.2964 ***	0.2945 ***	0.2980 ***
	(0.0437)	(0.0433)	(0.0435)	(0.0403)
Spline4×High Wage	0.2121 ***	0.2412 ***	0.2394 ***	0.2430 ***
	(0.0404)	(0.0387)	(0.0431)	(0.0394)
Spline5×High Wage	0.0696 ***	0.1086 ***	0.1102 ***	0.1182 ***
	(0.0403)	(0.0359)	(0.0486)	(0.0460)
Spline6×High Wage	-0.1886 ***	-0.1640 ***	-0.1655 ***	-0.1495 ***
	(0.0705)	(0.0550)	(0.0700)	(0.0681)
Spline1×Very High Wage	0.2026 **	0.2146 ***	0.2171 ***	0.2220 ***
	(0.1989)	(0.2017)	(0.2019)	(0.1815)
Spline2×Very High Wage	0.2724 ***	0.2866 ***	0.2822 ***	0.2897 ***
	(0.2182)	(0.2207)	(0.2216)	(0.1955)
Spline3×Very High Wage	0.1386 ***	0.1576 ***	0.1548 ***	0.1598 ***
	(0.0695)	(0.0716)	(0.0703)	(0.0627)
Spline4×Very High Wage	0.1639 ***	0.1855 ***	0.1832 ***	0.1871 ***
	(0.0405)	(0.0414)	(0.0413)	(0.0391)
Spline5×Very High Wage	0.1285 ***	0.1594 ***	0.1603 ***	0.1638 ***
	(0.0345)	(0.0290)	(0.0399)	(0.0365)
Spline6×Very High Wage	-0.1091 ***	-0.0821 ***	-0.0807 ***	-0.0718 ***
	(0.0366)	(0.0250)	(0.0471)	(0.0488)

(continued)

Variables	Without RE (1)	Normal RE (2)	Two Mass Points (3)	Three Mass Points (4)
First Birth:				
Center	-0.3596 ***	-0.4117 ***	-0.4097 ***	-0.4255 ***
	(0.0592)	(0.0675)	(0.0711)	(0.0690)
North	-0.4158 ***	-0.4768 ***	-0.4749 ***	-0.4916 ***
	(0.0718)	(0.0627)	(0.0830)	(0.0800)
Center×Wage	0.0481 ***	0.0270	0.0258	0.0187
	(0.0400)	(0.0438)	(0.0448)	(0.0440)
North×Wage	0.0535 ***	0.0325 *	0.0316 *	0.0236
	(0.3489)	(0.0357)	(0.4012)	(0.4581)
Second Birth:				
Constant	-3.1986 ***	-3.8702 ***	-3.5305 ***	-3.9537 ***
	(0.4966)	(0.2944)	(0.7169)	(1.0917)
Spline1×Low Wage	1.1041 ***	1.1333 ***	1.1256 ***	1.1385 ***
	(0.1276)	(0.0992)	(0.1261)	(0.1216)
Spline2×Low Wage	0.2915 ***	0.3490 ***	0.3393 ***	0.3494 ***
	(0.0698)	(0.0654)	(0.0715)	(0.0698)
Spline3×Low Wage	-0.0861 ***	-0.0359 *	-0.0309	-0.0349 *
	(0.0597)	(0.0522)	(0.0634)	(0.0614)
Spline4×Low Wage	-0.3980 ***	-0.3894 ***	-0.3860 ***	-0.3855 ***
	(0.1835)	(0.1193)	(0.1870)	(0.1830)
Spline1×Medium Wage	1.0263 ***	1.0611 ***	1.0518 ***	1.0652 ***
	(0.1178)	(0.1059)	(0.1191)	(0.1137)
Spline2×Medium Wage	0.3301 ***	0.3681 ***	0.3601 ***	0.3692 ***
	(0.0899)	(0.0716)	(0.0886)	(0.0856)
Spline3×Medium Wage	-0.1178 ***	-0.0805 ***	-0.0799 ***	-0.0813 ***
	(0.0444)	(0.0449)	(0.0480)	(0.0456)
Spline4×Medium Wage	-0.2553 ***	-0.2353 ***	-0.2324 ***	-0.2334 ***
	(0.1602)	(0.0772)	(0.1613)	(0.1582)
Spline1×High Wage	1.1208 ***	1.1516 ***	1.1406 ***	1.1569 ***
	(0.1358)	(0.1049)	(0.1368)	(0.1338)
Spline2×High Wage	0.2470 ***	0.2823 ***	0.2767 ***	0.2813 ***
	(0.0829)	(0.0758)	(0.0851)	(0.0807)
Spline3×High Wage	-0.1618 ***	-0.1328 ***	-0.1343 ***	-0.1354 ***
	(0.0392)	(0.0419)	(0.0425)	(0.0382)
Spline4×High Wage	-0.2158 ***	-0.2002 ***	-0.1998 ***	-0.1996 ***
	(0.1599)	(0.0558)	(0.1611)	(0.1628)
Spline1×Very High Wage	1.2812 ***	1.3095 ***	1.2999 ***	1.3154 ***
	(0.1485)	(0.0920)	(0.1485)	(0.1444)
Spline2×Very High Wage	0.1738 ***	0.2120 ***	0.2081 ***	0.2089 ***
	(0.0785)	(0.0652)	(0.0806)	(0.0743)
Spline3×Very High Wage	-0.2556 ***	-0.2304 ***	-0.2310 ***	-0.2328 ***
	(0.0475)	(0.0472)	(0.0489)	(0.0452)
Spline4×Very High Wage	-0.1606 ***	-0.1437 ***	-0.1436 ***	-0.1428 ***
	(0.0600)	(0.0599)	(0.0618)	(0.0548)

(continued)

Variables	Without RE (1)	Normal RE (2)	Two Mass Points (3)	Three Mass Points (4)
Second Birth:				
Age at the First Birth	-0.0315 ***	-0.0131 ***	-0.0137 ***	-0.0123 ***
	(0.0383)	(0.0092)	(0.0531)	(0.0551)
Center	-0.4483 ***	-0.5417 ***	-0.5409 ***	-0.5420 ***
	(0.0672)	(0.0828)	(0.0806)	(0.0756)
North	-0.4764 ***	-0.5835 ***	-0.5792 ***	-0.5820 ***
	(0.0734)	(0.0750)	(0.0857)	(0.0818)
Center×Wage	-0.0671 **	-0.0828 **	-0.0903 ***	-0.0819 **
	(0.0692)	(0.0744)	(0.0757)	(0.0702)
North×Wage	0.1191 ***	0.1257 ***	0.1161 ***	0.1233 ***
	(0.1183)	(0.0757)	(0.1589)	(0.2595)
Third Birth:				
Constant	-1.5729 ***	-2.4444 ***	-1.8571 ***	-2.6211 ***
	(0.6133)	(0.6630)	(0.6861)	(0.9666)
Spline1×Low Wage	0.7320 ***	0.7510 ***	0.7439 ***	0.7533 ***
	(0.2366)	(0.2367)	(0.2376)	(0.2224)
Spline2×Low Wage	0.0400	0.0489	0.0405	0.0577
	(0.1283)	(0.1322)	(0.1296)	(0.1149)
Spline3×Low Wage	-0.0535 ***	-0.0414 **	-0.0470 ***	-0.0378 **
	(0.0857)	(0.0362)	(0.0873)	(0.0870)
Spline1×Medium Wage	0.6577 ***	0.6714 ***	0.6638 ***	0.6780 ***
	(0.2550)	(0.2465)	(0.2558)	(0.2418)
Spline2×Medium Wage	0.0786	0.0796	0.0763	0.0838
	(0.1646)	(0.1652)	(0.1660)	(0.1518)
Spline3×Medium Wage	-0.1054 ***	-0.0965 ***	-0.1007 ***	-0.0932 ***
	(0.1116)	(0.0455)	(0.1111)	(0.1111)
Spline1×High Wage	0.7745 ***	0.7739 ***	0.7700 ***	0.7821 ***
	(0.2325)	(0.2133)	(0.2331)	(0.2214)
Spline2×High Wage	-0.0705	-0.0680	-0.0704	-0.0642
	(0.1582)	(0.1607)	(0.1601)	(0.1521)
Spline3×High Wage	-0.1037 ***	-0.0942 ***	-0.0981 ***	-0.0921 ***
	(0.0967)	(0.0506)	(0.0968)	(0.0961)
Spline1×Very High Wage	0.8597 ***	0.8575 ***	0.8556 ***	0.8638 ***
	(0.2663)	(0.2577)	(0.2690)	(0.2457)
Spline2×Very High Wage	-0.0752	-0.0749	-0.0762	-0.0726
	(0.1594)	(0.1601)	(0.1594)	(0.1499)
Spline3×Very High Wage	-0.1550 ***	-0.1473 ***	-0.1507 ***	-0.1449 ***
	(0.0650)	(0.0667)	(0.0672)	(0.0655)

(continued)

Variables	Without RE (1)	Normal RE (2)	Two Mass Points (3)	Three Mass Points (4)
Third Birth:				
Age at the Second Birth	-0.1161 ***	-0.0972 ***	-0.1066 ***	-0.0932 ***
	(0.0299)	(0.0188)	(0.0404)	(0.0512)
Center	-0.2898 ***	-0.3864 ***	-0.3422 ***	-0.4067 ***
	(0.2035)	(0.2129)	(0.2199)	(0.1986)
North	-0.2171 ***	-0.3057 ***	-0.2616 ***	-0.3293 ***
	(0.1758)	(0.1789)	(0.1889)	(0.1764)
Center×Wage	0.0495	0.0487	0.0460	0.0580
	(0.2068)	(0.2108)	(0.2090)	(0.1922)
North×Wage	0.1853 ***	0.1898 ***	0.1806 ***	0.2056 ***
	(0.1668)	(0.1740)	(0.1748)	(0.1626)
Random Effect:				
Normal Random Effect		0.5348 ***		
		(0.0774)		
Point1			-0.7600	-0.7794
Point2			0.2487 ***	0.3672 **
Point3				1.7051 ***
Weight1			0.54	0.33
Weight2			0.46	0.64
Weight3				0.03
No. Of Women	34,439	34,439	34,439	34,439
Log-Likelihood	-141666.01	-141623.93	-141627.93	-141619.68

Notes: RE= Random Effect, Two(Three) Mass Points indicate a finite mixture distribution for the Random Effect with two (three) mass points. Bootstrapped standard errors in parentheses. Standard errors are computed using 100 bootstrap replications of the 20% sample of the women aged between 15 to 40 in the ISAT-LFS, 2003. Spline First Birth: Spline1 [age 15-18), Spline2 [age 18-21), Spline3 [age 21-24), Spline4 [age 24-27), Spline5 [age 27-31), Spline6 [age 31-40). Spline Second Birth (Duration from the First Birth): Spline1 [Duration 0-2), Spline2 [Duration 2-4), Spline3 [Duration 4-8), Spline4 [Duration 8-26). Spline Third Birth (Duration from the Second Birth): Spline1 [duration 0-2), Spline2 [Duration 2-4), Spline3 [Duration 4-26). Reference Group for Region is South. Because we inserted an intercept in the model, when estimating a non parametric maximum likelihood with aML one masspoint has to be fixed (we fixed the first one).

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Chapter 3

The (mis)specification of Discrete Time Duration Models with Unobserved Heterogeneity: a Monte-Carlo study.

3.1 Introduction

Duration models are usually specified as hazard functions conditioned on a set of individual explanatory variables. One of the first and most common hazard functions used for continuous time durations is the proportional hazard, also called Cox model (see Cox 1972 and Cox and Oakes 1984). The discrete time analogue of the proportional hazard is given by the complementary log-log hazard model, i.e. a sequential binary model with error distributed as a Type-I extreme value distribution (for more details see Holford 1976; Prentice and Gloeckler 1978; Allison 1982; Narendranathan and Stewart 1993; Sueyoshi 1995). This model can be directly derived from the discretisation of a continuous time proportional hazard model;

it is therefore adequate to model continuous durations when observations are grouped into intervals. When instead the risk event, defining the end of a duration, can occur only in discrete time, then the duration cannot be considered as a continuous variable. In this case, more general sequential binary models can be considered, such as sequential probit and logit models.

One of the main issues concerning the estimation of both continuous and discrete time hazard regression models is the potential presence of unobserved individual characteristics, say unobserved heterogeneity (see Lancaster 1979, 1990; Heckman and Singer 1984; van den Berg 2001). Ignoring unobserved individual characteristics may bias estimates of the effect of observed explanatory variables and of duration dependence in the hazard function.

The standard way of accounting for unobserved heterogeneity is to consider a random component, which represents a scalar function of time-invariant unobserved variables. The hazard function is then defined conditioning on observed explanatory variables and on the unobserved random component. It is then possible to estimate model parameters by maximizing the likelihood from which the unobserved random effect is integrated out. The resulting model is a mixture of hazard functions with respect to the unobserved random component. The estimation of these mixture models requires either to assume a specific parametric distribution for the random component, or to use a non-parametric maximum likelihood estimation.

In principle the non-parametric maximum likelihood estimation is the best solution to minimize the potential bias caused by improper parametric distributional assumptions (see, for continuous time, Heckman and Singer 1984 and, for discrete time, Baker and Melino 2000; Zhang 2003). Nevertheless, the computation of the non-parametric estimator is not usually possible using commands built into common software packages. For this reason many non-specialists adopt easier estimation methods by either imposing specific parametric distributions for the unobserved heterogeneity or by ignoring altogether the unobserved heterogeneity.

In continuous time duration models, the unobserved heterogeneity distribution is often chosen to be gamma for analytical convenience (see Lancaster 1979) and theoretical reasons (see van den Berg 2001; and Abbring and van den Berg 2006). In discrete time duration models, instead, the assumption of a normal distribution can be computationally convenient. Under this assumption, discrete time duration models can be easily estimated as binary models with normal random effects using widely available statistical programs. The fact that discrete time duration models can easily estimated by using widely available statistical programs was first noticed by Yamaguchi (1991) and Jenkins (1995). Stata provides, for example, the commands `xtcloglog`, `xtlogit` and `xtprobit` (`cloglog`, `logit` and `probit`) to estimate binary models with normal random effects (without normal random effects) and error terms with extreme value, logistic and normal distributions.

In this chapter we evaluate the consequences of ignoring the unobserved heterogeneity or misspecifying its parametric distribution when estimating single spell discrete time duration models. Similar studies have been already carried out by Baker and Melino (2000), Zhang (2003), Gaure, Røed and Zhang (2005) (for continuous time duration models we refer instead to Heckman and Singer 1984; Lancaster 1985; Trussell and Richards 1985; Ridder 1987; Dolton and van der Klaauw 1995). They find that estimates are biased if unobserved heterogeneity is ignored or if its distribution is estimated non-parametrically using a discrete distribution but with an incorrect number of support points. Zhang (2003) and Gaure, Røed and Zhang (2005) find that the estimation bias is smaller if time-varying instead than time-invariant explanatory variables are used.

One important issue - overlooked by Baker and Melino (2000), Zhang (2003) and Gaure, Røed and Zhang (2005) - is that the residual variance in sequential binary models changes if unobserved heterogeneity is ignored or if a non-parametric distribution with too few or to many support points is used. Since the coefficients in binary models are usually normalized by dividing them by the residual standard deviation, models with high (low) residual variances produce coefficients which are attenuated (amplified). The attenuation (amplification) biases, Baker and Melino (2000), Zhang (2003) and Gaure, Røed and Zhang (2005) find, could be then due, at least in part, to their neglect of this issue. Mroz and Zaytas (2005) show that this is the case for the choice of the number of support points of the unobserved heterogeneity discrete distribution. Increasing (decreasing) the number of support points causes a reduction (rise) of the residual variance and an amplification (attenuation) of the coefficients. In this chapter we show that the attenuation bias caused by the omission of the unobserved heterogeneity or the misspecification of its distribution can be also a consequence, at least in part, of the coefficient normalization issue.

We undertake a Monte Carlo study to evaluate the effects of two misspecification problems in sequential logit models with unobserved heterogeneity:

1. omission of the unobserved heterogeneity when using time-varying and/or time-invariant explanatory variables,

2. imposing a normal distribution for the unobserved heterogeneity when instead the true distribution is a gamma or a discrete distribution.

The consequences of the first type of misspecification were already studied by Baker and Melino (2000) but they only considered time-invariant explanatory variables and, as already noted, without taking account of the normalization issue in binary models. The consequences of incorrectly imposing a normal random effect in discrete time duration model have not been studied before.

Finally, we consider the effect of misspecifying both the distribution of the unobserved heterogeneity and of the residual error.

The chapter is organized as follows. Section 3.2 considers the effects of neglecting unobserved heterogeneity while Section 3.3 considers the effects of its misspecification. In both sections we first discuss the theoretical consequences of omitting or misspecifying the unobserved heterogeneity, and we then assess those possible consequences through a Monte Carlo simulation exercise. In Section 3.4 we summarize the main findings.

3.2 Ignoring unobserved heterogeneity

3.2.1 Consequences of ignoring unobserved heterogeneity

Ignoring unobserved heterogeneity in duration models can cause a bias in the estimation of the duration dependence. More precisely, the omission of the unobserved heterogeneity causes an overestimation of the negative duration dependence (see for example Lancaster 1990; and van den Berg 2001). This is because people who have a high unobserved random component are more likely to experience the event of interest early, so that the sample of individuals that survive is a selected sample with relatively small random effects[1]. This selection process is known as *weeding out* or *sorting effect*.

Omitting unobserved heterogeneity may also bias the coefficients of the explanatory variables in the hazard model. For example neglecting unobserved heterogeneity in mixed proportional (continuous time) hazard models causes an underestimation of the proportionate response of the hazard function with respect to the explanatory variables (see van den Berg 2001 for a formal proof).

The bias is again due to a weeding out effect. Let us assume that the unobserved heterogeneity is given by a time-invariant scalar random effect, θ, independent of the explanatory variables; while the observed heterogeneity is given by a scalar function $\mu = m(X; \beta)$, where X is a vector of individual time-invariant explanatory variables and β is the vector of the corresponding coefficients. Without loss of generality, we assume in this section that the hazard function conditional on the observed explanatory variables and the unobserved heterogeneity be positively related to both θ and μ. A hazard model ignoring the unobserved heterogeneity is a hazard function conditional on the observed characteristics, X, but unconditional on the unobserved heterogeneity, θ, which we call the observed hazard function. The difference in the observed hazard function between survived people with high and low values of μ reflects also a gap in their values of θ. Survivors with a large μ have on average a smaller θ than survived people with a small μ, so that the difference between the observed hazard functions is on average lower than the difference we would observe if the survivors had the same value for θ. If we fail to recognize that the lower difference between the observed hazards is due to

[1]Notice that, without loss of generality, we are assuming in this section that the unobserved random component be positively related to the hazard function.

a difference in the unobserved heterogeneity, we would erroneously estimate an attenuated effect of the explanatory variables on the hazard.

More rigorously, the weeding out effect on the covariate coefficients can be described as the consequence of a lack of independence between the random effect for a generic individual i, θ_i, and her (his or its) observed heterogeneity, $m(X_i;\beta)$, given a duration $T_i \geq \tau$, where τ is a scalar strictly higher than zero, say the failure of the condition $(\theta_i \perp\!\!\!\perp m(X_i;\beta) \mid T_i \geq \tau)$. Notice, instead, that hazard models assume that $(\theta_i \perp\!\!\!\perp X_i)$ which implies that $(\theta_i \perp\!\!\!\perp m(X_i;\beta) \mid T_i \geq 0)$. We assume here that $(T_i \mid X_i, \theta_i)$ be identically and independently distributed i.i.d. across individuals.

There are some continuous time duration models for which the attenuation bias due to omitted unobserved heterogeneity reduces to a rescaling by a factor (a bias proportionally identical) for all explanatory variables coefficients or to a bias only for the intercept. Lancaster (1985) proves analytically that the omission of unobserved heterogeneity in mixed proportional hazard models with baseline distribution given by a Weibull causes a rescaling by a constant factor for all coefficients. Ridder (1987) proves analytically that the omission in mixed proportional hazard models with known baseline hazard and with no right censuring causes a bias only for the intercept. Moreover, Ridder (1987) suggests that replacing the baseline with a non-parametric flexible specification should produce an almost unbiased estimation of the covariates coefficients.

Ridder's suggestion is supported by his Monte Carlo study and by some other empirical studies: see Dolton and van der Klaauw (1995); Meyer (1990); Trussell and Richards (1985). By contrast, the conjecture is not confirmed by the Monte Carlo experiment in Baker and

Melino (2000), who consider discrete time duration models with single spell. But this contradictory result may be due to the fact that in discrete time duration models the coefficients are identified only up to a scale normalization and models with different specifications use different normalizations, which Baker and Melino (2000) do not consider.

It is possible to prove analytically that the omission of the unobserved heterogeneity causes only a rescaling by a factor of the covariate coefficients when considering sequential probit models with normal random effects θ_{it} that are i.i.d. across individuals and time t, and independent of the explanatory variables, X_{it}, and with known duration dependence function. This is because $(\theta_{it} \perp\!\!\!\perp m(X_{it}; \beta) \mid T_i \geq \tau)$ for any $\tau \geq 0$. (see Appendix B for more details).

Similar analytical results do not exist instead for more general discrete time duration models. In this chapter we consider the consequences of omitting unobserved heterogeneity in more general single spell discrete time duration models. In particular we consider the following cases:

a. the unobserved random effects is time-invariant and follows a normal, a gamma or a discrete distribution with two points of support,

b. the error distribution is logistic instead of normal,

c. the duration dependence is ignored or it is approximated by a flexible function,

d. the covariates are i.i.d. across individuals and time or i.i.d. across individuals but not time.

Cases described in (a) to (c) were already considered by Baker and Melino (2000). They find that ignoring the unobserved heterogeneity component causes an attenuation bias for the covariate coefficients. In this chapter we replicate their Monte Carlo study to evaluate again the consequences of ignoring unobserved heterogeneity but taking into account the issue of the coefficient normalization. Mroz and Zayats (2005) reconsider instead the Monte Carlo study of Baker and Melino (2000) to compare the effects of alternative non-parametric specifications of the unobserved heterogeneity distribution when taking account of the normalization issue. Baker and Melino (2000) find that non-parametric maximum likelihood estimation that penalizes specifications with many mass points for the unobserved heterogeneity distribution produces more reliable coefficients; but Mroz and Zayats (2005) give evidence that this result is a consequence of the normalization issue.

Case (d) is an extension necessary to understand how the estimation bias can depend on the types of covariates used. If the covariates, say X_{it} for individual i and duration (time) t, are i.i.d. across individuals and time, then estimation bias should reduce. This is because in this case the independence between the unobserved component and the observed covariates holds even when conditioning on survival until a time strictly greater than zero, that is $(\theta_i \perp\!\!\!\perp X_{it} \mid T_i \geq \tau)$, where $\tau > 0$.

If, instead, covariates are i.i.d. across individuals but time-invariant or correlated across time, then we would expect an attenuation bias. Nevertheless, the bias could consist in a proportional reduction, in absolute value, in all covariate coefficients, i.e. a rescaling by a constant factor.

In next section we describe the Monte Carlo experiment carried out in order to study the potential consequences of omitting time-invariant unobserved random effects for the cases described by (a) to (d).

3.2.2 Description of the Monte Carlo Simulation: Data Generating Processes

We consider the same data generating processes (DGPs from now on) used in the Monte Carlo study of Baker and Melino (2000) except that we use both time-varying and time-invariant explanatory variables while they only use a time-invariant one.

We assume that duration is measured in discrete time. This is quite often the case when observations are grouped into intervals or when the event, whose occurrence defines the end of a duration, can occur only in discrete time. We then record an event taking place in the interval $(t-1, t]$ as occurred in t.

We assume that the probability of experiencing an event in t (or in the time interval $(t-1, t]$) conditional on survival to $(t-1)$ for a generic individual i is given by :

$$Pr(d_{it} = 1 | d_{it-1} = 0) \quad = \quad Pr(z^*_{it} < 0 | z^*_{it-1} \geq 0) \tag{3.2.1}$$

where d_{it} is a dummy variable indicating the event occurrence at t for individual i, and z^*_{it} is a latent continuous variable which is lower than zero if $d_{it} = 1$ and higher or equal to zero otherwise. We assume that z^*_{it} obeys the following linear model:

$$z^*_{it} = X_{it}\beta - f(t) + \theta_i + \epsilon_{it} \tag{3.2.2}$$

where X_{it} is a vector of explanatory variables, β is the corresponding vector of parameters, $f(t)$ is a deterministic function of elapsed duration, θ_i is an individual random effect representing unobserved heterogeneity, ϵ_{it} is a residual error term distributed as a logistic with zero mean and variance $\pi^2/3$ and both θ_i and ϵ_{it} are independent of the explanatory variables[2]. We can then write the hazard probability conditional on the observed explanatory

[2]The definition of the above discrete time hazard model and the notation used are consistent with Baker

variables, X_{it}, and on the unobserved heterogeneity, θ_i, as

$$Pr(d_{it} = 1 | d_{it-1} = 0, X_{it}, \theta_i) = \frac{1}{1 + exp(z_{it})} \tag{3.2.3}$$

where

$$z_{it} = X_{it}\beta - f(t) + \theta_i. \tag{3.2.4}$$

By choosing different specifications for the observed explanatory variables, X_{it}, the duration dependence function, $f(t)$, and the unobserved heterogeneity, θ_i, we produce a set of different DGPs.

We organize the simulations in two main sets. In the first set, exercise **A**, we focus on the effect of omitting unobserved heterogeneity when using different types of explanatory variables. In particular we consider three DGPs using three different typologies of observed explanatory variables: **A1** time-varying variables, **A2** time-invariant variables and **A3** variables given by the sum of a time-invariant variable and a time-varying one, say mixture variables. For each choice of the covariates we consider two different types of duration dependence function, one increasing and one deceasing, and three types of distribution for the unobserved heterogeneity, a discrete (with two points of support), a gamma and a normal distribution. This provides us with 18 different DGPs.

In the second set of simulations, exercise **B**, we consider both time-invariant and time-varying covariates and we focus on the effect of omitting the unobserved heterogeneity when considering or not considering duration dependence in the simulated and estimated models. Again we consider three different types of distribution of the unobserved heterogeneity,

and Melino (2000) except for the negative sign in front of the duration dependence function $f(t)$. This is because it is counterintuitive to have a negative (positive) duration dependence when $f(t)$ is increasing (decreasing) in time.

whereas we consider only one specification for the duration function and for the vector of covariates which includes both time-invariant and mixture variables. This second simulation exercise produces six different types of DGPs.

For each of the DGPs in simulation exercise **A** we consider two sample sizes: 500 and 1000 individuals. For simulation exercise **B**, we consider instead three sample sizes: 500, 1000 and 5000 individuals. The higher sample size of 5000 is motivated by the fact that in exercise **B** there are some small sample biases which decrease very slowly with the sample size.

As in Baker and Melino (2000) we draw 100 samples for each DGP, we follow the individuals for 40 periods and consider all durations greater than 40 as censored.

In the following, we discuss in more detail how we specify the explanatory variables, the duration dependence function and the unobserved heterogeneity distribution for different types of DGP.

Observed explanatory variables.

As in Baker and Melino (2000) we fix the variance of the observed heterogeneity in the hazard model, $Var(X_{it}\beta)$, to be equal to 0.25 for all our simulations.

In exercise **A** we specify the observed heterogeneity in the hazard model as:

$$X_{it}\beta \quad = \quad X_{1,\,it}\beta_1 + X_{2,\,it}\beta_2. \tag{3.2.5}$$

where $X_{1,it}$ and $X_{2,it}$ are normal random variables, and β_1 and β_2 are fixed parameters which we set to be equal to 1 and 0.5.

We consider three different simulations for the variables, $X_{1,it}$ and $X_{2,it}$:

A1 two independent time-varying variables identically and independently distributed (i.i.d.)

across individuals and time with zero means and variances 0.125 and 0.5;

A2 two independent time-invariant variables i.i.d. across individuals with zero means and variances 0.125 and 0.5;

A3 two independent variables defined as the sum of a time-invariant variable and a time-varying one with equal variances, say mixture variables; more precisely, $X_{1,it}$ ($X_{2,it}$) is the sum of a time-varying variable defined as **A1** but with variance 0.0625 (0.25) and a time-invariant variable defined as in **A2** but with variance 0.0625 (0.25).

Simulation **A1** represents an extreme case which is interesting from a theoretical point of view but it is less interesting from an empirical one. In empirical examples explanatory variables are usually correlated across time so that the assumption of explanatory variables i.i.d. across individuals and time does not seem to be very plausible. Simulation **A2** represents the opposite extreme case where all the explanatory variables are supposed to be time-invariant. This is the case considered by Baker and Melino (2000). Finally, simulation **A3** represents an intermediate case where the explanatory variables are given by the sum of a time-invariant component and a time-varying one. Earnings and income can be examples of such types of variables. Earnings and income (or their logarithm transformations) are usually assumed by economists to be the sum of a permanent component and a transitory one (see for example Moffitt and Gottschalk, 2002).

In simulation exercise **B** we specify instead the observed heterogeneity in the hazard model as:

$$X_{it}\,\beta \;=\; X_{1,\,i}\,\beta_1 + X_{2,\,i}\,\beta_2 + X_{1,\,it}\,\beta_3 + X_{2,\,it}\,\beta_4 \qquad (3.2.6)$$

where $X_{1,i}$ and $X_{2,i}$ are time-invariant variables, $X_{3,it}$ and $X_{4,it}$ are mixture variables and $\beta' = [1, 0.5, 1, 0.5]$. To be more specific $X_{1,i}$ and $X_{2,i}$ are time-invariant variables defined as in **A2** but with variances 0.0625 and 0.25, $X_{3,it}$ and $X_{4,it}$ are mixture variables defined as in **A3** but with variances 0.0625 and 0.25, and all explanatory variables are independent.

Duration Dependence.

In exercise **A** we consider, as in Baker and Melino (2000), the following deterministic time function

$$f(t) = 1 - \exp\left(\frac{1-t}{5}\right) \tag{3.2.7}$$

for a negative duration dependence and

$$f(t) = \exp\left(\frac{1-t}{5}\right) - 1 \tag{3.2.8}$$

for a positive duration dependence.

In simulation exercise **B** we consider instead $f(t) = 0$ for no duration dependence and again $f(t) = \exp\left(\frac{1-t}{5}\right) - 1$ for a positive duration dependence.

Unobserved Heterogeneity.

In both exercises **A** and **B** we consider three distributions for the unobserved heterogeneity θ_i: discrete, gamma and normal distribution. To be consistent with Baker and Melino (2000) we set $E(\theta_i) = 1.8$ and $Var(\theta_i) = 1$ and for the discrete distribution we consider two support points with equal probability, that is:

$$\theta_i = \begin{cases} 0.8 & \text{with probability } 0.5 \\ 2.8 & \text{with probability } 0.5\,. \end{cases} \tag{3.2.9}$$

3.2.3 Description of the Monte Carlo Simulation: estimation models

Using the data simulated in exercise **A** we estimate a sequential logit model as specified in (3.2.3) but ignoring the unobserved heterogeneity and approximating the duration dependence function with either a cubic polynomial in t or using a 'non-parametric' step function. As in Baker and Melino (2000) we consider a step function given by

$$\phi(t) = \sum_{\tau=1}^{40} \phi_\tau D_{t\tau} \tag{3.2.10}$$

where

$$D_{t\tau} = \begin{cases} 1 & \text{if } t = \tau \\ 0 & \text{otherwise} \end{cases}$$

and ϕ_τ, $\tau = 1, ..., 40$, are the corresponding coefficients. However, because few individuals survive after 15 periods, we allow the coefficients to vary for each period only until $\tau = 14$ and then we impose constant coefficients within the following time intervals: $\tau = 15 - 19$, $20 - 24$, $25 - 29$, $30 - 40$.

Using the data simulated in exercise (**B**) we estimate again a sequential logit model ignoring the unobserved heterogeneity and approximating the duration dependence function with either a zero function (no duration dependence) or the above 'non-parametric' step function.

3.2.4 Results

In this section we present the results of the Monte Carlo simulation exercises **A** and **B**.

The results of the exercise **A** are reported in Table 3.1, which is divided in three panels providing the results for time-varying covariates (top panel **A1**), time-invariant covariates (middle panel **A2**) and mixture covariates (bottom panel **A3**). The results reported are the average and the standard deviation over 100 replications for the covariate coefficients, β_1 (the true value of which is 1) and β_2 (which true value is 0.5), and their ratio β_1/β_2. By row we specify the type of DGP used to generate the simulated data. More precisely, we consider six different types of DGPs: sequential logit model with negative or positive duration dependence and with unobserved heterogeneity following a discrete, a gamma or a normal distribution (labeled discrete, gamma and normal). By column we specify instead the sample size (500 or 1000 observations/individuals) and the type of estimation model used: sequential logit model omitting random effects and with duration dependence approximated by a step function (labeled Step DD) or by a cubic polynomial (labeled polynomial DD).

If the omission of the unobserved heterogeneity causes only a rescaling by a factor of the coefficients, then the coefficients would be biased towards zero (attenuation bias) but their ratio would still be correctly estimated. This seems supported by the results in Table 3.1 when using any type of covariates. Moreover, when using time-varying covariates which are i.i.d. across individuals and time (top panel **A1**), the attenuation problem for the coefficients does not seem to be significant. When using instead covariates which are i.i.d. across individuals and time-invariant, the attenuation problem is more severe. Finally, when using mixture covariates the attenuation bias magnitude seems to be intermediate between the two previous extreme cases.

Using different distributions for the simulated unobserved heterogeneity components and different specifications for the simulated duration dependence (negative or positive) produce some very small and insignificant differences in the coefficients.

Approximating the duration dependence function by using a step or a cubic polynomial function does not produce any significant difference in the results.

Finally, increasing the sample size from 500 to 1000 observations leads to a slight improvement in the results, meaning that the attenuation bias for β_1 and β_2 decreases a little and the average ratio between coefficients becomes even closer to the true value of two.

To summarize, ignoring unobserved heterogeneity in sequential logit models seems to cause an attenuation of the covariate coefficients due to a different normalization. This attenuation bias cancels almost completely when using covariates which are i.i.d. across individuals and time, while it does not when the covariates are highly autocorrelated.

As emphasized in Section 3.2.1, ignoring unobserved heterogeneity may cause a bias for the covariate coefficients as well as for the duration dependence function estimation. In Figures 3.1 and 3.2 we compare the true simulated duration dependence function (line labeled true duration) with the estimated duration dependence functions averaged over 100 samples of size 1000, simulated using the different DGPs in Monte Carlo exercise **A1**. We consider a negative true duration dependence function given by (3.2.7) in Figure 3.1 and a positive one given by (3.2.8) in Figure 3.2. In both Figures the three average estimated dependence functions (lines labeled discrete, gamma and normal) are obtained by considering a cubic polynomial in the duration and by using data simulated from three different DGPs assuming a discrete, a gamma or a normal distribution for the unobserved random effects.

It seems that ignoring the unobserved heterogeneity causes an overestimation of the negative duration dependence and a spurious negative dependence even when the true duration dependence is positive.

In Tables 3.2 and 3.3 we report the results for exercise **B** in which we consider a hazard model with two time-invariant and two time-varying covariates. The estimation models considered are sequential logit models ignoring unobserved heterogeneity. In Table 3.2 the duration dependence is estimated using a step function, whereas in Table 3.3 the duration dependence is ignored. In both Tables we report the average and the standard deviation over 100 replications for the covariate coefficients, $\beta' = [\beta_1, \beta_2, \beta_3, \beta_4]$ which have true values $[1, 0.5, 1, 0.5]$, and the ratios between any pair of coefficients. The estimation results for different simulated DGPs are reported by column. We consider six different DGPs, sequential logit models with three possible distributions for the unobserved heterogeneity (discrete, gamma or normal) and with either negative or no duration dependence. Both Tables are divided in 3 panels corresponding to three different sample sizes (500, 1000 and 5000).[3]

In Table 3.2 the covariate coefficients seem to be significantly underestimated. Moreover, the underestimation of the coefficients seems to be slightly larger for the pair of time-invariant variables than for the pair of time-varying ones. In other words, it seems that the rescaling factor is slightly dissimilar for different types of variables (time-varying and invariant variables). Indeed, ratios between coefficients seems to be correctly estimated when considering two variables of the same type (see β_1/β_2 and β_3/β_4) and to be slightly biased when considering the ratio between two different types of variables (see β_3/β_2, β_3/β_1, β_1/β_4 and β_4/β_3). Nevertheless, since the standard deviations for coefficient ratios are quite high, the differences in the rescaling factor are not significant. This result is confirmed even when using a larger sample size of 5000 observations. In conclusion, we find again that omitting the

[3] As already said, the higher sample size of 5000 is motivated by the fact that in exercise **B** there are some small sample biases which decrease very slowly with the sample size.

unobserved heterogeneity cause an attenuation of the covariate coefficients due to a rescaling factor which differs slightly and not significantly by typology of variable.

Looking at the results in Table 3.3 where the estimation models ignore both the unobserved heterogeneity and the duration dependence, the underestimation of covariate coefficients reduces and the rescaling factor is more similar for variables of different types. The ratios between coefficients are not biased especially when considering a sample size of 5000 observations. This result is not unexpected because the unobserved heterogeneity component and the time duration component are negatively and positively related to the duration, so that the biases for ignoring those components should go in opposite directions and should offset each other, at least in part.

In conclusion, the two main findings of this section are that ignoring the unobserved heterogeneity in sequential logit models causes an overestimation of the negative duration dependence and an attenuation of the covariate coefficients but this attenuation is due to a rescaling by a factor of every coefficient by the same amount. Since coefficients in binary models are only identified up to a scale normalization, inference will not be affected by the unobserved heterogeneity omission except for the duration dependence.

3.3 Misspecifying the unobserved heterogeneity distribution

3.3.1 Consequences of misspecifying unobserved heterogeneity

Heckman and Singer (1984) argue that an incorrect assumption about the distribution of the unobserved heterogeneity in hazard models can have severe consequences. In particular, they find that the parameter estimates for a model with Weibull baseline hazard are very sensitive to changes in the distribution assumed for the unobserved heterogeneity. Similar results were found also by Trussell and Richards (1985), Hougaard, Myglegaard and Borch-Johnsen (1994), Baker and Melino (2000), Zhang (2003) and Gaure, Røed and Zhang (2005). However, Ridder and Verbakel (1983) criticize the findings of Heckman and Singer (1984) and highlight the fact that a non-flexible specification of the baseline hazard may explain their (Heckman and Singer 1984) findings.

We reconsider the heterogeneity misspecification problem in the specific case of discrete time duration models with single spell specified as sequential binary models. We are particularly interested in evaluating the effect of imposing a normal distribution for the unobserved heterogeneity component when the true distribution is a gamma or a discrete distribution with two support points.

By contrast, Baker and Melino (2000) studied the effect of using different non parametric specification of the unobserved heterogeneity distribution. They find that if too many support points for the estimated heterogeneity distribution are used, the unobserved heterogeneity dispersion is overestimated and the covariate coefficients are biased away from zero (amplification bias). As explained by Mroz and Zayats (2005) this amplification bias may simply be due to a rescaling by a factor of the variables. Mroz and Zayats (2005) find indeed that the effects of the covariates seem to be better estimated if the number of support points is large when taking into account the normalization problem.

In addition, we consider the potential consequences of misspecifying the distribution of the residual error as well as of the unobserved heterogeneity in the sequential binary models.

3.3.2 Description of the Monte Carlo simulation: DGPs and estimation models

As in Section 3.2, we carry out a Monte Carlo experiment by simulating 100 samples from a set of different DGPs (data generator processes).

The DGPs used to generate the data are sequential logit models with unobserved heterogeneity following three alternative types of distribution (discrete, gamma or normal), with a negative time duration dependence and two explanatory variables given by two mixture variables. For more details on the DGPs we refer to Monte Carlo exercise **A3** described in Section 3.2.2.

Our estimation models are instead given by sequential binary models with normal random effects and duration dependence approximated by a cubic polynomial in the duration. We consider three models: (1) sequential logit, (2) sequential probit and (3) sequential complementary log-log models. We estimate those sequential binary models with random effects by using Stata which considers an adaptive Gauss-Hermite quadrature to approximate the integral of the maximum likelihood function with respect to the random effects (see for more details StataCorp 2005). An alternative estimation methods is given by the simulated maximum likelihood, see for more details Gourieroux and Monfort (1996) and Train (2003).

The simulation exercise is carried out as the previous ones by drawing 100 samples for each DGP and three different sample sizes: 500, 1000 and 5000 individuals. We consider durations longer than 40 periods as censored.

3.3.3 Results

In Tables 3.4, 3.5 and 3.6 we report the results corresponding to the use of the three different estimation models: sequential logit, probit and or complementary log-log with normal random effects and cubic polynomial in the duration. The simulated data used in all three Tables are generated from the same DGP: a sequential logit model with negative duration dependence and unobserved heterogeneity following three alternative distributions (discrete, gamma or normal).

In each Table we report the average and the standard deviation over 100 replications for the two covariate (mixture variable) coefficients, β_1 (which true value is 1) and β_2 (which true value is 0.5), their ratio β_1/β_2, the fraction of residual variance explained by individual random effects (ρ), the average number of iterations and the number of cases out of 100 of successful convergence of the maximum likelihood algorithm. We report averages and standard deviations only for the cases where convergence was reached. Each Table is divided in three panels reporting results produced using three different sample sizes: 500, 1000 and 5000 observations.

Looking at the results in Table 3.4, where both estimation and simulated models are sequential logit models, the covariate coefficients do not seem to be underestimated. They seem to be well estimated even when the unobserved heterogeneity distribution is misspecified. This is an encouraging result for practitioners who would like to use easy-to-implement estimation methods to take account of unobserved heterogeneity.

In Table 3.5, where the estimation model is given by a sequential probit model while the true DGPs are given by sequential logit models, the two covariate coefficients are underestimated but the ratio between them is still unbiased. Again we do not find relevant differences when considering DGPs with different distributions for the random effects.

Since the logistic distribution is similar to the normal one but with heavier tails, the difference in the estimated coefficients when using probit instead of logit models is probably due to the different normalization implied by the different residual variances. Because the residual variance in logit models is normalized to $\pi^2/3$ while in probit models is normalized to one, the rescaling factor should be given by $\pi/\sqrt{3}$ (see Greene 2003) . By multiplying the coefficients in Table 3.5 by this factor, we find that the estimates are very close to the ones reported in Table 3.4.

Finally, in Table 3.6, we change the estimation model to a sequential complementary log-log model. The two covariate coefficients seem to be slightly underestimated while the ratio between them is unbiased. The coefficients seem slightly lower than the ones shown in Table 3.4 and it seems that coefficients bias be due again to a rescaling. Again, the results do not seem to be affected by the distribution assumed for the unobserved heterogeneity in the DGPs.

Increasing the sample size has the same effect for all three types of models (logit, probit and complementary log-log): the attenuation bias does not change significantly, the standard deviations decrease, and the number of unsuccessful convergence cases reduces to zero.

The fraction of the residual variance explained by the individual unobserved heterogeneity, ρ, seems very slightly and insignificantly underestimated when using sequential logit models. It is still slightly and insignificantly underestimated when using sequential complementary log-log models, and it is more significantly underestimated when considering a sequential probit. Notice that a higher underestimation of the ρ coefficient seems to be associated with a higher attenuation bias for the coefficients. This result seems to confirm Baker and Melino's (2000) conclusion that an underestimation (overestimation) of the dispersion of the unobserved heterogeneity leads to an attenuation (amplification) of the covariate coefficients. However, we find that this attenuation (amplification) bias is simply due to a normalization problem that Baker and Melino (2000) did not notice.

To evaluate the effect of misspecifying the unobserved heterogeneity distribution on the duration dependence estimation, we plot the baseline hazard functions estimated using sequential logit, probit and complementary log-log with normal random effects: see Figures 3.3, 3.4 and 3.5. In all three Figures we consider data simulated from a DGP given by a sequential logit model with the usual three types of distribution for the unobserved heterogeneity.

When both estimation and simulation models are given by a sequential logit model (Figure 3.3), the true baseline hazard (simulated) has a profile similar to the three estimated baseline hazards averaged over 100 samples (labeled discrete, gamma and normal) corresponding to three different DGPs, i.e. sequential logit models with random effects following a discrete, a gamma and a normal distribution.

When we change the estimation model to a sequential probit model with normal random effects (Figure 3.4), the estimated baseline hazards (labeled discrete, gamma and normal) have instead a different profile with respect to the true baseline hazard (labeled simulated) especially for long durations.

Finally, when using a sequential complementary log-log model with normal random effect for the estimation of the duration model, we find that the profile of the estimated baseline hazards (labeled discrete, gamma and normal in Figure 3.5) follow the true one (simulated), but the negative dependence is overestimated for short durations.

In summary, it seems that misspecification of the unobserved heterogeneity distribution does not seriously affect the estimation results. Changes in the error distribution (logistic, normal and extreme value) bias the duration dependence estimation but cause only a rescaling of the coefficients estimates by a constant factor. Because coefficients in binary models are identified only up to a scale normalization, the rescaling is not a genuine problem.

3.4 Conclusions

This chapter assesses the effects of ignoring unobserved heterogeneity or misspecifying its distribution in single spell discrete time duration models. In particular, we focus on assessing the consequences of adopting two models that can be easily estimated using standard software: sequential binary models with or without individual normal random effects.

The main findings from our Monte Carlo study can be summarized as follows. First, neglecting the unobserved heterogeneity seems to cause a bias in the duration dependence estimation. It does not seem to cause a bias in the covariate coefficients but rather a rescaling by a constant factor. Second, the rescaling factor is close to one when considering covariates i.i.d. across individuals and time, while it is significantly smaller than one for covariates that are i.i.d. across individuals and correlated across time. Third, misspecifying the random effects distribution biases neither the duration dependence nor the covariate coefficients estimation. Fourth, misspecifying the error distribution, assuming a normal or an extreme value distribution instead than a logistic one, seems to cause a bias in the duration dependence estimation while it seems to cause only a proportional rescaling of the covariate coefficients.

These findings are very encouraging for practitioners who estimate discrete time duration models with or without normal random effects using command built into standard statistical software packages.

Table 3.1: Means and standard deviations of the coefficients estimates over 100 samples. Monte Carlo exercise A.

	500 Observations						*1000 Observations*					
	Step DD			*Polynomial DD*			*Step DD*			*Polynomial DD*		
DGP	β_1	β_2	β_1/β_2	β_1	β_2	β_1/β_2	β_1	β_2	β_1/β_2	β_1	β_2	β_1/β_2
True Value	1	0.5	2	1	0.5	2	1	0.5	2	1	0.5	2
*Time-varying covariates, **A1***												
Positive duration dependence												
Discrete UH	0.896	0.446	2.061	0.892	0.445	2.058	0.934	0.473	1.999	0.932	0.472	1.998
(sd)	0.151	0.073	0.479	0.148	0.072	0.471	0.092	0.052	0.302	0.091	0.051	0.300
Gamma UH	0.967	0.464	2.144	0.963	0.462	2.146	0.951	0.478	2.015	0.948	0.477	2.015
(sd)	0.147	0.074	0.505	0.146	0.074	0.505	0.086	0.052	0.300	0.085	0.051	0.299
Normal UH	0.922	0.459	2.081	0.919	0.457	2.082	0.928	0.474	1.978	0.926	0.473	1.979
(sd)	0.158	0.087	0.526	0.158	0.086	0.526	0.099	0.050	0.296	0.099	0.049	0.295
Negative duration dependence												
Discrete UH	0.903	0.444	2.086	0.900	0.442	2.090	0.911	0.458	2.014	0.909	0.457	2.015
(sd)	0.142	0.070	0.472	0.142	0.069	0.469	0.090	0.053	0.308	0.090	0.053	0.307
Gamma UH	0.927	0.454	2.106	0.922	0.450	2.109	0.911	0.456	2.023	0.908	0.455	2.021
(sd)	0.152	0.074	0.531	0.148	0.072	0.525	0.107	0.057	0.306	0.107	0.056	0.306
Normal UH	0.895	0.446	2.047	0.891	0.443	2.051	0.905	0.464	1.977	0.903	0.463	1.977
(sd)	0.167	0.073	0.459	0.165	0.072	0.457	0.095	0.053	0.319	0.095	0.052	0.321
*Time-invariant covariates, **A2***												
Positive duration dependence												
Discrete UH	0.662	0.326	2.168	0.664	0.326	2.168	0.678	0.315	2.250	0.680	0.316	2.248
(sd)	0.147	0.077	0.790	0.147	0.077	0.787	0.112	0.060	0.675	0.112	0.060	0.672
Gamma UH	0.672	0.334	2.121	0.673	0.334	2.122	0.659	0.328	2.082	0.660	0.329	2.081
(sd)	0.144	0.075	0.688	0.145	0.075	0.687	0.104	0.061	0.533	0.104	0.061	0.530
Normal UH	0.730	0.350	2.277	0.731	0.350	2.274	0.684	0.341	2.051	0.685	0.341	2.051
(sd)	0.141	0.084	1.167	0.140	0.084	1.148	0.102	0.050	0.432	0.101	0.050	0.432
Negative duration dependence												
Discrete UH	0.740	0.355	2.161	0.739	0.354	2.167	0.726	0.351	2.113	0.724	0.350	2.113
(sd)	0.141	0.070	0.551	0.142	0.070	0.559	0.102	0.051	0.435	0.102	0.051	0.437
Gamma UH	0.625	0.316	2.075	0.622	0.314	2.075	0.613	0.303	2.105	0.611	0.303	2.105
(sd)	0.132	0.067	0.654	0.133	0.066	0.660	0.099	0.058	0.588	0.099	0.058	0.588
Normal UH	0.709	0.342	2.194	0.707	0.341	2.192	0.660	0.340	1.985	0.659	0.339	1.986
(sd)	0.136	0.082	0.681	0.136	0.081	0.677	0.102	0.052	0.427	0.102	0.052	0.428
*Mixture covariates, **A3***												
Positive duration dependence												
Discrete UH	0.789	0.390	2.114	0.788	0.390	2.117	0.809	0.399	2.068	0.809	0.399	2.069
(sd)	0.154	0.083	0.636	0.153	0.083	0.638	0.113	0.058	0.418	0.114	0.058	0.419
Gamma UH	0.810	0.408	2.077	0.807	0.407	2.077	0.794	0.404	2.005	0.793	0.403	2.007
(sd)	0.159	0.082	0.637	0.159	0.083	0.635	0.115	0.056	0.414	0.114	0.056	0.414
Normal UH	0.811	0.402	2.087	0.808	0.401	2.086	0.810	0.402	2.056	0.810	0.401	2.057
(sd)	0.148	0.073	0.546	0.148	0.073	0.548	0.105	0.058	0.386	0.106	0.058	0.388
Negative duration dependence												
Discrete UH	0.828	0.410	2.082	0.824	0.408	2.084	0.835	0.416	2.044	0.833	0.414	2.048
(sd)	0.157	0.079	0.534	0.157	0.078	0.537	0.101	0.054	0.391	0.102	0.054	0.398
Gamma UH	0.791	0.390	2.110	0.787	0.387	2.113	0.761	0.375	2.074	0.759	0.374	2.073
(sd)	0.141	0.074	0.598	0.141	0.074	0.593	0.109	0.055	0.419	0.108	0.055	0.422
Normal UH	0.796	0.393	2.083	0.792	0.391	2.086	0.786	0.393	2.025	0.785	0.392	2.027
(sd)	0.153	0.067	0.553	0.153	0.067	0.557	0.102	0.053	0.313	0.102	0.053	0.313

Note: Characteristics of the DGPs (data generator processes) and of the estimation models are given by row and by column. UH = unobserved heterogeneity. DD = duration dependence. Step = step function. Poynomial= cubic polynomial function.

Table 3.2: Means and standard deviations of the coefficients estimates over 100 samples. Monte Carlo exercise B. Estimation model with step function duration dependence.

	β_1	β_2	β_3	β_4	β_1/β_2	β_3/β_4	β_3/β_2	β_3/β_1	β_1/β_4	β_4/β_2
True Value	1	0.5	1	0.5	2	2	2	1	2	1
DGP										
					500 Observations:					
Positive duration dependence:										
Discrete UH	0.479	0.251	0.575	0.273	2.091	2.327	2.512	1.377	1.940	1.194
	(0.155)	(0.070)	(0.131)	(0.077)	(0.994)	(0.995)	(1.024)	(0.699)	(1.035)	(0.564)
Gamma UH	0.507	0.236	0.584	0.295	2.470	2.128	2.886	1.260	1.867	1.456
	(0.148)	(0.077)	(0.150)	(0.074)	(1.428)	(0.821)	(1.628)	(0.560)	(0.846)	(0.966)
Normal UH	0.489	0.239	0.574	0.283	2.664	2.225	3.237	1.295	1.863	1.531
	(0.145)	(0.080)	(0.170)	(0.071)	(3.397)	(1.116)	(5.616)	(0.601)	(0.858)	(2.068)
No duration dependence:										
Discrete UH	0.510	0.255	0.589	0.283	2.144	2.197	2.517	1.258	1.897	1.197
	(0.139)	(0.064)	(0.146)	(0.061)	(0.857)	(0.772)	(1.074)	(0.504)	(0.707)	(0.489)
Gamma UH	0.453	0.239	0.548	0.275	2.115	2.153	2.596	1.428	1.812	1.270
	(0.144)	(0.073)	(0.149)	(0.066)	(1.030)	(0.928)	(1.360)	(1.256)	(1.012)	(0.546)
Normal UH	0.471	0.235	0.553	0.278	2.276	2.119	2.640	1.348	1.792	1.319
	(0.148)	(0.072)	(0.144)	(0.067)	(1.242)	(0.808)	(1.738)	(0.734)	(0.758)	(0.647)
					1000 Observations:					
Positive duration dependence:										
Discrete UH	0.469	0.233	0.557	0.276	2.226	2.109	2.680	1.259	1.770	1.306
	(0.105)	(0.061)	(0.115)	(0.054)	(1.093)	(0.674)	(1.637)	(0.423)	(0.545)	(0.628)
Gamma UH	0.469	0.238	0.554	0.282	2.131	2.034	2.503	1.231	1.720	1.277
	(0.096)	(0.055)	(0.093)	(0.052)	(0.943)	(0.508)	(0.976)	(0.331)	(0.483)	(0.484)
Normal UH	0.498	0.239	0.566	0.284	2.216	2.053	2.518	1.197	1.806	1.272
	(0.101)	(0.054)	(0.112)	(0.050)	(0.799)	(0.542)	(0.856)	(0.389)	(0.498)	(0.462)
No duration dependence:										
Discrete UH	0.496	0.248	0.576	0.284	2.130	2.097	2.480	1.209	1.800	1.219
	(0.100)	(0.056)	(0.108)	(0.054)	(0.881)	(0.535)	(1.008)	(0.340)	(0.468)	(0.445)
Gamma UH	0.444	0.224	0.547	0.271	2.122	2.106	2.588	1.316	1.699	1.288
	(0.104)	(0.051)	(0.100)	(0.053)	(0.864)	(0.616)	(0.859)	(0.462)	(0.516)	(0.475)
Normal UH	0.487	0.233	0.552	0.286	2.259	1.995	2.563	1.194	1.754	1.327
	(0.106)	(0.057)	(0.100)	(0.050)	(0.934)	(0.518)	(0.982)	(0.351)	(0.514)	(0.514)
					5000 Observations:					
Positive duration dependence:										
Discrete UH	0.472	0.231	0.561	0.278	2.070	2.033	2.463	1.198	1.708	1.219
	(0.045)	(0.027)	(0.045)	(0.023)	(0.289)	(0.266)	(0.354)	(0.157)	(0.210)	(0.154)
Gamma UH	0.466	0.236	0.562	0.285	1.989	1.987	2.405	1.221	1.643	1.217
	(0.047)	(0.024)	(0.045)	(0.022)	(0.289)	(0.236)	(0.334)	(0.168)	(0.212)	(0.152)
Normal UH	0.481	0.243	0.568	0.283	2.004	2.019	2.362	1.192	1.709	1.177
	(0.049)	(0.023)	(0.047)	(0.024)	(0.303)	(0.233)	(0.300)	(0.144)	(0.222)	(0.144)
No duration dependence:										
Discrete UH	0.502	0.245	0.570	0.285	2.067	2.013	2.349	1.143	1.773	1.172
	(0.040)	(0.024)	(0.044)	(0.023)	(0.258)	(0.226)	(0.306)	(0.129)	(0.207)	(0.127)
Gamma UH	0.445	0.223	0.549	-0.279	2.014	1.979	2.490	1.248	1.603	1.264
	(0.046)	(0.025)	(0.043)	(0.023)	(0.303)	(0.218)	(0.348)	(0.161)	(0.213)	(0.162)
Normal UH	0.467	0.235	0.566	-0.279	2.006	2.039	2.431	1.224	1.682	1.198
	(0.047)	(0.023)	(0.046)	(0.020)	(0.285)	(0.225)	(0.304)	(0.150)	(0.207)	(0.140)

Note: Characteristics of the DGPs (data generator processes) are given by row. UH = unobserved heterogeneity.

Table 3.3: Means and standard deviations of the coefficients estimates over 100 samples. Monte Carlo exercise B. Estimation model ignoring duration dependence.

	β_1	β_2	β_3	β_4	β_1/β_2	β_3/β_4	β_3/β_2	β_3/β_1	β_1/β_4	β_4/β_2
True Value	1	0.5	1	0.5	2	2	2	1	2	1
DGP										
500 Observations:										
Positive duration dependence:										
Discrete UH	0.627	0.330	0.636	0.297	2.099	2.412	2.140	1.189	2.393	1.001
	(0.216)	(0.094)	(0.155)	(0.088)	(1.054)	(1.184)	(0.985)	(0.657)	1.539)	(0.495)
Gamma UH	0.624	0.310	0.632	0.327	2.231	2.187	2.327	1.124	2.101	1.182
	(0.186)	(0.097)	(0.187)	(0.094)	(0.996)	(1.247)	(1.261)	(0.545)	(0.974)	(0.576)
Normal UH	0.638	0.311	0.635	0.312	2.676	2.266	2.722	1.108	2.234	1.314
	(0.198)	(0.109)	(0.196)	(0.086)	(3.082)	(1.219)	(3.943)	(0.525)	(1.109)	(1.688)
No duration dependence:										
Discrete UH	0.657	0.329	0.654	0.311	2.133	2.234	2.165	1.091	2.239	1.019
	(0.184)	(0.082)	(0.169)	(0.070)	(0.840)	(0.839)	(0.954)	(0.448)	(0.914)	(0.425)
Gamma UH	0.548	0.277	0.611	0.294	2.371	2.490	2.454	1.315	2.214	1.311
	(0.194)	(0.106)	(0.192)	(0.091)	(2.619)	(2.113)	(3.618)	(0.883)	(1.680)	(2.051)
Normal UH	0.599	0.301	0.611	0.307	2.275	2.159	2.302	1.193	2.092	1.146
	(0.195)	(0.095)	(0.165)	(0.081)	(1.262)	(0.918)	(1.532)	(0.709)	(0.941)	(0.594)
1000 Observations:										
Positive duration dependence:										
Discrete UH	0.614	0.305	0.608	0.303	2.264	2.102	2.308	1.054	2.119	1.118
	(0.141)	(0.083)	(0.133)	(0.061)	(1.220)	(0.732)	(1.786)	(0.372)	(0.715)	(0.620)
Gamma UH	0.620	0.299	0.641	0.314	2.232	2.101	2.345	1.103	2.032	1.151
	(0.149)	(0.078)	(0.133)	(0.058)	(0.860)	(0.544)	(1.153)	(0.394)	(0.582)	(0.530)
Normal UH	0.652	0.309	0.624	0.316	2.240	2.047	2.150	1.013	2.135	1.094
	(0.138)	(0.072)	(0.132)	(0.058)	(0.806)	(0.589)	(0.740)	(0.339)	(0.609)	(0.400)
No duration dependence:										
Discrete UH	0.638	0.318	0.635	0.316	2.133	2.085	2.138	1.041	2.089	1.058
	(0.131)	(0.072)	(0.122)	(0.063)	(0.878)	(0.560)	(0.917)	(0.308)	(0.573)	(0.403)
Gamma UH	0.590	0.278	0.579	0.296	2.279	2.050	2.212	1.059	2.081	1.138
	(0.151)	(0.067)	(0.122)	(0.063)	(0.895)	(0.636)	(0.766)	(0.409)	(0.700)	(0.402)
Normal UH	0.622	0.295	0.606	0.318	2.277	1.970	2.215	1.030	2.013	1.167
	(0.140)	(0.074)	(0.115)	(0.057)	(0.934)	(0.555)	(0.833)	(0.317)	(0.606)	(0.445)
5000 Observations:										
Positive duration dependence:										
Discrete UH	0.615	0.302	0.616	0.306	2.064	2.035	2.072	1.012	2.024	1.025
	(0.060)	(0.036)	(0.054)	(0.027)	(0.292)	(0.288)	(0.314)	(0.142)	(0.247)	(0.133)
Gamma UH	0.615	0.301	0.627	0.310	2.064	2.037	2.108	1.033	1.997	1.043
	(0.068)	(0.033)	(0.049)	(0.027)	(0.326)	(0.250)	(0.298)	(0.146)	(0.298)	(0.149)
Normal UH	0.626	0.317	0.631	0.312	1.997	2.036	2.009	1.018	2.021	0.994
	(0.065)	(0.031)	(0.058)	(0.028)	(0.310)	(0.248)	(0.273)	(0.131)	(0.272)	(0.130)
No duration dependence										
Discrete UH	0.645	0.316	0.632	0.316	2.060	2.012	2.021	0.987	2.053	1.009
	(0.053)	(0.031)	(0.052)	(0.026)	(0.255)	(0.234)	(0.265)	(0.119)	(0.244)	(0.114)
Gamma UH	0.552	0.287	0.603	0.299	1.956	2.034	2.134	1.107	1.862	1.056
	(0.068)	(0.034)	(0.057)	(0.028)	(0.370)	(0.281)	(0.358)	(0.160)	(0.293)	(0.153)
Normal UH	0.597	0.301	0.629	0.307	2.005	2.059	2.112	1.064	1.953	1.032
	(0.061)	(0.030)	(0.055)	(0.024)	(0.291)	(0.248)	(0.278)	(0.135)	(0.253)	(0.128)

Note: Characteristics of the DGPs (data generator processes) are given by row. UH = unobserved heterogeneity.

Table 3.4: Means and standard deviations of coefficient estimates over 100 samples. Estimation model: sequential logit. DGP: sequential logit.

	β_1	β_2	β_1/β_2	$\rho = \frac{\sigma_\theta^2}{\sigma_\epsilon^2+\sigma_\theta^2}$	Iterations	Convergence
True Value	1	0.5	2	$\frac{1}{\frac{\pi^2}{3}+1} = 0.233$		
DGP						
			500 Observations:			
Negative duration dependence:						
Discrete Unobserved Het.	0.927	0.455	2.130	0.174	6.908	98
	(0.170)	(0.093)	(0.611)	(0.143)	(2.981)	
Gamma Unobserved Het.	0.913	0.452	2.110	0.118	5.404	99
	(0.183)	(0.087)	(0.672)	(0.072)	(2.263)	
Normal Unobserved Het.	0.923	0.460	2.083	0.148	6.271	96
	(0.158)	(0.086)	(0.550)	(0.116)	(2.759)	
			1000 Observations:			
Negative duration dependence:						
Discrete Unobserved Het.	0.937	0.462	2.069	0.156	6.424	99
	(0.137)	(0.071)	(0.405)	(0.093)	(2.607)	
Gamma Unobserved Het.	0.942	0.474	2.012	0.151	5.889	99
	(0.123)	(0.060)	(0.317)	(0.090)	(2.788)	
Normal Unobserved Het.	0.944	0.470	2.048	0.170	6.316	98
	(0.121)	(0.066)	(0.393)	(0.118)	(2.775)	
			5000 Observations:			
Negative duration dependence:						
Discrete Unobserved Het.	0.956	0.477	2.013	0.202	6.850	100
	(0.079)	(0.040)	(0.178)	(0.099)	(2.851)	
Gamma Unobserved Het.	0.916	0.458	2.003	0.110	4.190	100
	(0.064)	(0.027)	(0.163)	(0.059)	(1.522)	
Normal Unobserved Het.	0.966	0.482	2.009	0.188	5.850	100
	(0.076)	(0.034)	(0.161)	(0.102)	(2.455)	

Note: Iterations = average number of iterations for the convergence of the likelihood maximization algorithm. Convergence = number of cases over 100 replications of successful convergence. ρ = fraction of residual variance explained by individual random effects.

Table 3.5: Means and standard deviations of coefficient estimates over 100 samples. Estimation model: sequential probit. DGP: sequential logit.

	β_1	β_2	β_1/β_2	$\rho = \frac{\sigma_\theta^2}{\sigma_\epsilon^2+\sigma_\theta^2}$	Iterations	Convergence
True Value	1	0.5	2	$\frac{1}{1+1} = 0.5$		
DGP						
500 Observations:						
Negative duration dependence:						
Discrete Unobserved Het.	0.519	0.255	2.128	0.306	8.690	100
	(0.099)	(0.058)	(0.587)	(0.140)	(1.495)	
Gamma Unobserved Het.	0.543	0.270	2.069	0.307	8.535	99
	(0.102)	(0.049)	(0.522)	(0.126)	(1.358)	
Normal Unobserved Het.	0.528	0.264	2.089	0.304	8.455	99
	(0.102)	(0.058)	(0.566)	(0.157)	(2.370)	
1000 Observations:						
Negative duration dependence:						
Discrete Unobserved Het.	0.525	0.257	2.078	0.292	8.848	99
	(0.082)	(0.039)	(0.407)	(0.104)	(1.480)	
Gamma Unobserved Het.	0.534	0.270	2.013	0.305	8.410	100
	(0.087)	(0.039)	(0.396)	(0.112)	(1.326)	
Normal Unobserved Het.	0.537	0.268	2.049	0.324	8.760	100
	(0.079)	(0.041)	(0.407)	(0.124)	(1.457)	
5000 Observations:						
Negative duration dependence:						
Discrete Unobserved Het.	0.524	0.261	2.015	0.311	9.010	100
	(0.036)	(0.017)	(0.174)	(0.044)	(1.259)	
Gamma Unobserved Het.	0.542	0.271	2.004	0.305	8.740	100
	(0.037)	(0.015)	(0.159)	(0.042)	(0.960)	
Normal Unobserved Het.	0.538	0.268	2.010	0.315	8.850	100
	(0.038)	(0.016)	(0.161)	(0.049)	(1.175)	

Note: Iterations = average number of iterations for the convergence of the likelihood maximization algorithm. Convergence = number of cases over 100 replications of successful convergence. ρ = fraction of residual variance explained by individual random effects.

Table 3.6: Means and standard deviations of coefficient estimates over 100 samples. Estimation model: sequential complementary log-log. DGP: sequential logit.

	β_1	β_2	β_1/β_2	$\rho = \frac{\sigma_\theta^2}{\sigma_\epsilon^2+\sigma_\theta^2}$	Iterations	Convergence
True Value	1	0.5	2	$\frac{1}{\frac{\pi^2}{6}+1} = 0.378$		
DGP						
			500 Observations:			
Negative duration dependence:						
Discrete Unobserved Het.	0.855	0.421	2.123	0.243	6.714	98
	(0.147)	(0.086)	(0.606)	(0.127)	(2.428)	
Gamma Unobserved Het.	0.861	0.442	2.010	0.222	5.890	100
	(0.156)	(0.072)	(0.545)	(0.112)	(2.238)	
Normal Unobserved Het.	0.859	0.430	2.076	0.224	6.358	95
	(0.152)	(0.081)	(0.556)	(0.130)	(2.475)	
			1000 Observations:			
Negative duration dependence:						
Discrete Unobserved Het.	0.871	0.429	2.069	0.233	6.602	98
	(0.123)	(0.062)	(0.410)	(0.094)	(2.560)	
Gamma Unobserved Het.	0.886	0.432	2.084	0.228	5.730	100
	(0.107)	(0.051)	(0.381)	(0.084)	(2.348)	
Normal Unobserved Het.	0.879	0.437	2.053	0.250	6.240	100
	(0.109)	(0.062)	(0.386)	(0.131)	(2.590)	
			5000 Observations:			
Negative duration dependence:						
Discrete Unobserved Het.	0.882	0.440	2.013	0.285	6.710	100
	(0.058)	(0.029)	(0.178)	(0.089)	(2.388)	
Gamma Unobserved Het.	0.878	0.438	2.011	0.224	5.310	100
	(0.054)	(0.028)	(0.160)	(0.075)	(2.246)	
Normal Unobserved Het.	0.901	0.450	2.008	0.287	6.290	100
	(0.060)	(0.027)	(0.161)	(0.097)	(2.262)	

Note: Iterations = average number of iterations for the convergence of the likelihood maximization algorithm. Convergence = number of cases over 100 replications of successful convergence. ρ = fraction of residual variance explained by individual random effects.

Figure 3.1: Estimated and true negative duration dependence functions. Monte Carlo exercise A1. Unobserved heterogeneity ignored.

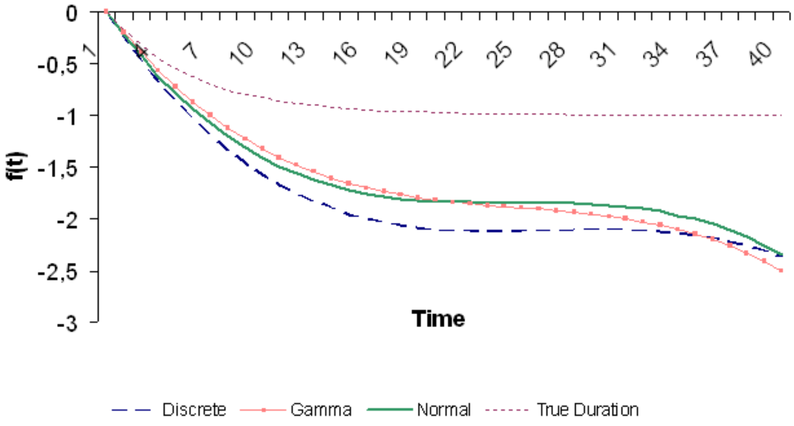


Figure 3.2: Estimated and true positive duration dependence functions. Monte Carlo exercise A1. Unobserved heterogeneity ignored.

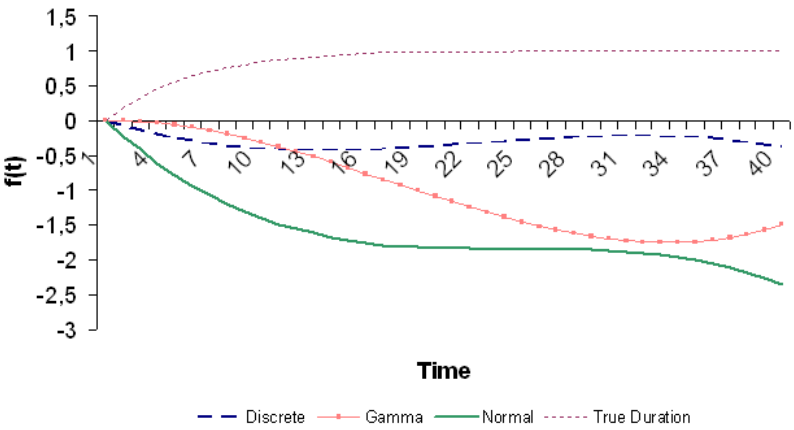


Figure 3.3: Estimated and true baseline hazards. Estimation model: sequential logit with normal random effects. DGP: sequential logit with unobserved heterogeneity.

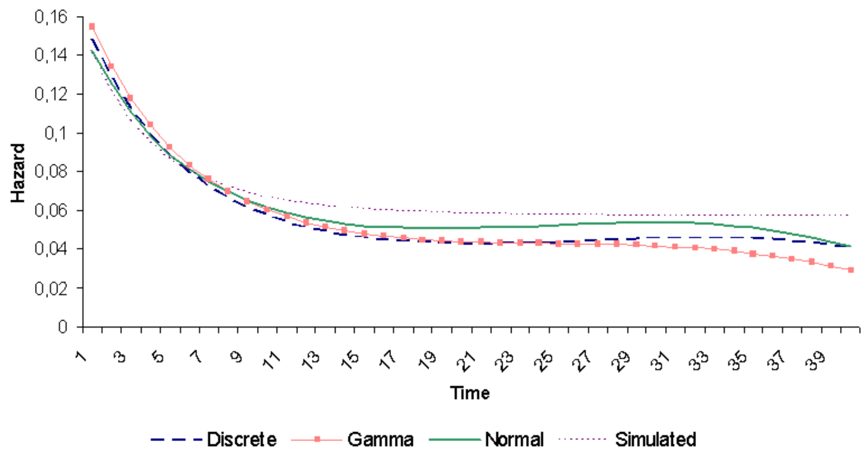


Figure 3.4: Estimated and true baseline hazards. Estimation model: sequential probit with normal random effects. DGP: sequential logit with unobserved heterogeneity.

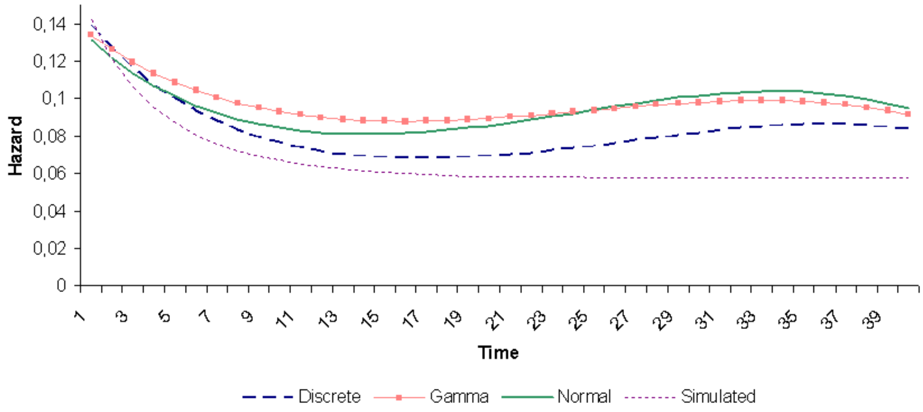
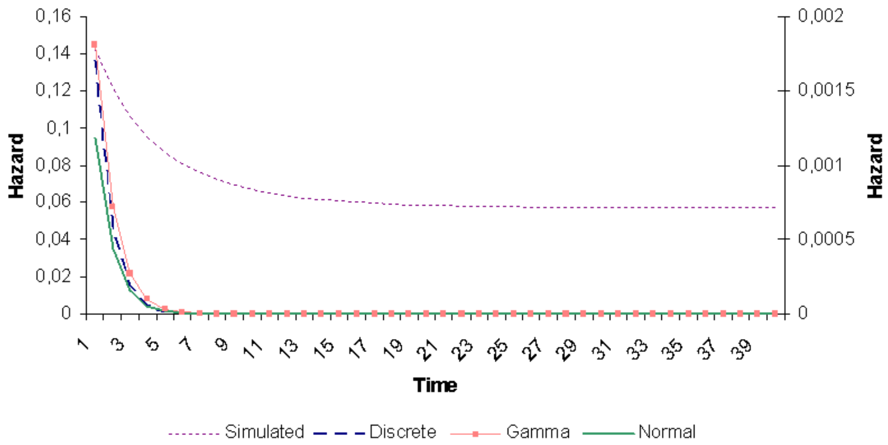


Figure 3.5: Estimated and true baseline hazards. Estimation model: sequential complementary log-log with normal random effects. DGP: sequential logit with unobserved heterogeneity.



Appendix A

Wage Equation

Wage equations are commonly estimated (see for instance Mincer 1974) by linear regression where the dependent variable is the natural log of the reported wage. Typical variables to be included are age and age squared, education (either in terms of years of education or as a dummy variable reflecting the educational level), type of education, work experience, number of children, age of the children, ethnicity, region, profession. In theory we can estimate separate wage equations for men and women within the status of working persons and pensioners. We do exclude pensioners: women in pensionable age are excluded as we limited our analysis to those who are 40 or less. Male are not considered because of the established unit of analysis. Husbands' wage is ignored: since we only know the marital status in 2003 (and information about the wedding date is not provided) we are prevented from reconstructing it retrospectively. However, there will be women who are recorded with zero wage simply because they do not work. The problem is that they might have chosen not to work because they would receive relatively low wage. In terms of the economic theory, they do not work because their offered wage is lower than their reservation wage (i.e. the lowest wage for which they would chose to work). If we choose those who work in our wage equation only, we do get a selection bias. One standard solution for this problem is to estimate a participation equation using a probit model (the so-called generalized Tobit or Tobit of second type, see Gourieroux 2000).

Consider the market wage (Y_{mi}) and the reservation wage (Y_{ri}) of person i specified as following:

$$Y_{mi} = x_i\prime\beta + \epsilon_i \tag{A.0.1}$$

and

$$Y_{ri} = z_i\prime\gamma + \nu_i \tag{A.0.2}$$

where $x_i\prime$ ($z_i\prime$) indicate the set of covariates, $\beta(\gamma)$ the corresponding parameters and ϵ_i (ν_i) the error terms for the market wage, (A.0.1), (reservation wage, (A.0.2)) equation.

The observed wage is given by:

$$Y_i = \begin{cases} Y_{mi} & \text{if } Y_{mi} \geq Y_{ri} \\ 0 & \text{Otherwise} \end{cases}$$

with:

$$Y_i = \begin{cases} x_i\prime\beta + \epsilon_i & \text{if } \epsilon_i - \nu_i \geq z_i\prime\gamma - x_i\prime\beta \\ 0 & \text{Otherwise} \end{cases} \tag{A.0.3}$$

The two step estimation provides a first step for the estimation of a probit model for the probability to be in the labour market and a second step for the estimation of the regression model with an additional variable (the inverse Mill's ratio) using the sub-sample of individual with $d_i = 1$, where:

$$d_i = \begin{cases} 1 \quad \text{if } Y_{mi} \geq Y_{ri}, \quad \text{i.e. a woman works} \\ 0 \quad \text{if } Y_{mi} < Y_{ri}, \quad \text{i.e. a woman does not work} \end{cases}$$

In our case, we are prevented from using this standard solution because of the combination of two different and independent data sets. Mapping the predicted BOI-SHIW wage into the ISTAT-LFS data set requires to have exactly the same variables in both data sets. Including information about the husbands as regressors (i.e. $z_i\prime$) in the participation equation attenuate the bias of the coefficients but we impute, to the women of the ISTAT-LFS data set, a predicted wage which captures year by year the presence of an additional wage even if we cannot address [1] the presence of an husband for that year.

Given this setting, the lack of data available and the data mapping we estimate a Tobit model (a Tobit model of the first type or Tobit with deterministic censure, see Tobin 1958) which censors the wage distribution at selected (upper or lower) point. The woman wage equation is left censored at zero.

Define:

$$\begin{cases} Y: & \text{Hourly Observed Wage} \Rightarrow \text{min (Y)=0;} \\ a: & \text{Positive Constant;} \\ W = Y + a: & a - \text{augmented Hourly Wage} \Rightarrow \text{min (W)=a;} \\ w = \lg W: & \Rightarrow \text{min (w)= lg } a. \end{cases}$$

Given this definition,

[1] Problem of retrospectively constructing the marital status.

$$Y_i = \begin{cases} Y_{mi} & \text{if } Y_{mi} \geq 0 \\ 0 & \text{if } Y_{mi} < 0. \end{cases} \tag{A.0.4}$$

We are interested in taking lg of (A.0.4), but the logarithm of zero is $-\infty$. Taking a positive constant a and assuming that $W_i = Y_i + a$, we get:

$$W_i = \begin{cases} a + x_i\prime\beta + \epsilon_i & \text{if } x_i\prime\beta \geq -\epsilon_i \\ a & \text{Otherwise} \end{cases}$$

with $\min(W_i) = a$ and:

$$w_i = \lg(W_i) = \begin{cases} \lg(a + x_i\prime\beta + \epsilon_i) & \text{if } x_i\prime\beta \geq -\epsilon_i \\ \lg(a) & \text{Otherwise} \end{cases}$$

and $\min(w_i) = \lg a$.

$\lg a$ is the lower bound for the log hourly wage distribution and it is the point of censure of the Tobit model. All the observations with augmented log hourly wage less than $\lg a$ are censored. If we apply to this model a standard OLS, the estimates are biased and inconsistent (Tobin, 1958).

If we denote $\hat{\omega}_{it}$ the log hourly predicted wage for the i-th woman at time t, the wage equation we estimate becomes:

$$\hat{\omega}_{it} = \hat{\beta}_0 + \hat{\beta}_1 \text{age}_{it} + \hat{\beta}_2 (\text{age})^2_{it} + \hat{\beta}_3 \sum_{j=1}^{8} \text{education}_{itj} + \hat{\beta}_4 \sum_{j=1}^{20} \text{region}_{itj} \tag{A.0.5}$$

where $t = 1983, ..., 2003$, age_{it} is the age of the i-th woman at time t, $(\text{education})_{itj}$, for $j = 1, ..., 8$, are eight dummy variables for different levels of education attained from the i-th woman at time t and $(\text{region})_{itj}$, for $j = 1, ..., 20$, are twenty dummy variables, one for each of the Italian regions. Table A.1 presents the estimated coefficient from an hourly wage equation for women in a Tobit model using the BOI-SHIW data set.

The selected sample in Table A.1 contains 4,749 women in the condition of being mothers, i.e. less than 45 years, not pensioners and not studying. 2,601 of these women are censored because they exhibit a zero wage.

As we expected, the coefficient for age is positive and for age squared is negative: this is in line with the economic/econometric literature that wage increase with age but at decreasing rate (Mincer 1974). Our results confirms the classical way to take education as a proxy of wage: the larger you study the more you earn (Schultz 1985). Moving from lower level of education to higher level induces a gain in terms of hourly wage rates. We do not include position in the woman wage equation estimated from BOI-SHIW because it is not provided the the ISTAT-LFS. To be more precise, the ISTAT-LFS Survey asks women position only at the time of the interview and we are no able to reconstruct it back in 1983. The hourly wage does not exhibit very large variability across different region of Italy, even if there is a clear tendency of lower hourly wage in the South (linked both with the presence of public job and the percentage of women not working) when compared to the North one.

Table A.1: Estimated Hourly Woman Wage Equation (Tobit Model)

Variable	Coefficient	(Standard Error)
Age:		
Age	0.276**	(0.021)
$(\text{Age})^2$	-0.003**	(0.000)
Education:		
Middle School	0.530**	(0.092)
Professional Secondary School Diploma	1.246**	(0.137)
High School	1.541**	(0.096)
Short Course University Degree	2.118**	(0.301)
Bachelor's Degree	2.458**	(0.124)
Post-Graduate Qualification	3.375**	(0.772)
Regions:		
Piemonte	1.744**	(0.171)
Valle d' Aosta	1.339*	(0.658)
Lombardia	1.477**	(0.165)
Trentino	2.099**	(0.258)
Veneto	1.321**	(0.181)
Friuli	1.553**	(0.229)
Liguria	1.675**	(0.196)
Emilia Romagna	1.881**	(0.170)
Toscana	1.608**	(0.181)
Umbria	1.782**	(0.208)
Marche	1.728**	(0.197)
Lazio	0.767**	(0.188)
Abruzzi	1.024**	(0.220)
Molise	0.967**	(0.320)
Campania	-0.143	(0.182)
Basilicata	1.466**	(0.325)
Calabria	1.107**	(0.230)
Sicilia	0.191	(0.179)
Sardegna	1.065**	(0.208)
Constant	-7.157**	(0.509)
No. of Women	4,749	
No. of censored Women	2,601	
Pseudo R^2	0.149	
Log-L	-5596.9167	

Notes: Dependent variable is the woman log hourly wage. Equation estimated from the BOI-SHIW, 2002. Reference group for education: none and elementary school. Reference group for regions: Puglia.

$^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$.

Appendix B

Probit Model

Let us consider a sequential probit model with unobserved heterogeneity given by a normal random variable identically and independently distributed (i.i.d.) across individuals and time. Moreover, let d_{it} be a dummy variable indicating the risk event occurrence for the generic individual i $(i = 1, ..., n)$ at time, duration, t $(t = 1, ..., T)$. Then the hazard probability for a generic $i-th$ individual at time t is given by

$$Pr(d_{it} = 1 | d_{it-1} = 0) \quad = \quad \Phi(X_{it}\beta + f(t) + \theta_{it}), \tag{B.0.1}$$

where X_{it} is a vector of explanatory variables, β is the vector of the corresponding coefficients, $f(t)$ is a known deterministic function of the duration, θ_{it} is an unobserved random variable i.i.d. across individuals and time following a normal distribution with zero mean and variance σ_θ^2 and independent of the explanatory variables, and Φ is the Gaussian cumulative distribution. If θ_{it} were observed and $f(t)$ were known, then the maximum likelihood estimation of the sequential probit conditioning to X_{it} and θ_{it} would produce consistent estimates for the β coefficients.

If we omit the unobserved heterogeneity, the hazard probability would be instead:

$$Pr(d_{it} = 1 | d_{it-1} = 0) \quad = \quad \Phi(X_{it}\tilde{\beta} + \tilde{f}(t)), \tag{B.0.2}$$

where $\tilde{\beta} = \frac{\beta}{\sigma}$, $\tilde{f}(\cdot)$ is given by $\frac{1}{\sigma}f(\cdot)$, and $\sigma = \sqrt{\sigma_\theta^2 + 1}$. Assuming that $\tilde{f}(\cdot)$ is known, we can still use the maximum likelihood method to estimate consistently the explanatory variable coefficients (see Maddala, 1987). Notice that the weeding out effect does not operate here because the unobserved heterogeneity is given by a random variables i.i.d. across individuals and time. People who survive at time t because of a low unobserved random component may have a high unobserved random component in $(t + 1)$, so that θ_{it} and X_{it} are independent conditioning to $t = 0$ but also conditioning to $t > 0$.

Finally, notice that the new coefficients in (B.0.2) are rescaled and therefore not directly comparable with the coefficients in the hazard model (B.0.1) (see Arulampalam, 1999). Since the rescaling factor is given by $\frac{1}{\sigma}$ and $\sigma > 1$, ignoring the unobserved heterogeneity causes an attenuation bias for the covariate coefficients.

Bibliography

Aassve A., Billari F.C., Mazzucco S. , Ongaro F. (2002) Leaving home: A comparative analysis of ECHP data, *Journal of European Social Policy*, 12, 4, 289–305.

Abbring J.H., van den Berg G.J. (2006) The unobserved heterogeneity distribution in duration analysis, Tinbergen Institute Discussion Papers 06-059/3, Tinbergen Institute.

Abramowitz M., Stegun I.A. (1972) *Handbook of Mathematical Functions.*, Dover Publications Inc., New York.

Allison P. (1982) Discrete-time methods for the analysis of event histories, in: *Sociological Methodology*, Leinhardt S., ed., Jossey-Bass Publishers, San Francisco, 61–97.

Amuedo-Dorantes C., Kimmel J. (2005) The motherhood wage gap for women in the united states: The importance of college and fertility delay, *Review of Economics of the Household*, 3, 1, 17–48.

Arellano M., Meghir C. (1992) Female labour supply and on-the-job search: An empirical model estimated using complementary data sets, *Review of Economic Studies*, 59, 3, 537–559.

Arumlampalam W. (1999) A note on estimated coefficients in random effects probit models, *Oxford Bulletin of Economics and Statistics*, 61, 4, 597–602.

Baker M., Melino A. (2000) Duration dependence and nonparametric heterogeneity: a Monte Carlo study, *Journal of Econometrics*, 96, 357–393.

Becker G.S. (1960) An economic analysis of fertility, in: *Demographic and Economic Change in Developed Countries*, Princeton University Press, Princeton, NJ, 209–231.

Becker G.S. (1965) A theory of the allocation of time, *Economic Journal*, 75, 493–517.

Becker G.S. (1981) *A Treatise on the Family*, Harvard University Press, Cambridge, MA.

Becker G.S. , Lewis H.G. (1973) On the interaction between the quantity and quality of children, *The Journal of Political Economy*, 81, 2, S279–S288.

Bernardi L. (2003) Channels of social influence on reproduction, *Population Research and Policy Review*, 22, 527–555.

Billari F.C., Borgoni R. (2004) Assessing the use of sample selection models in the estimation of fertility postponement effects, *Statistical Methods and Applications*, 14, 3, 389 – 402.

Billari F.C. , Kohler H.P. (2002) The impact of union formation dynamics on first births in West Germany and Italy: are there signs of convergence?, in: *Dynamics of Fertility and Partnership in Europe: Insights and Lessons from Comparative Research, Volume 2*, Klijzing E., Corijn M., eds., United Nations, Geneva/New York, 43–58.

Billari F.C. , Kohler H.P. (2004) Patterns of Low and Lowest Low Fertility in Europe, *Population Studies*, 58, 2, 161–176.

Blossfeld H.P. , De Rose A. (1992) Educational expansion and the changes in entry into marriage and motherhood: the experience of Italian women, *Genus*, 48, 3-4, 73–89.

Boix C. (1997) Mercado de trabajo y declive demográphico, *El Pais*, July 3.

Brewster K.L. , Rindfuss R.R. (2000) Fertility and women's employment in industrialized nations, *Annual Review of Sociology*, 26, 271–296.

Bumpass L.L. , Mburugu E.K. (1977) Age at marriage and completed family size, *Social Biology*, 24, Spring, 31–37.

Bumpass L.L., Rindfuss R.R. , Janosik R.B. (1978) Age and marital status at first birth and the pace of subsequent fertility marriage and completed family size, *Demography*, 15, 1, 75–86.

Butz W. , Ward M. (1979) The emergence of countercyclical U.S. fertility, *American Economic Review*, 69, 3, 318–328.

Cigno A. (1991) *Economics of the Family*, Clarendon Press, Oxford.

Cigno A. (1994) Economic considerations in the timing of births: Theory and Evidence, in: *The Family, the Market and the State in Ageing Societies*, Ermisch J. , Ogawa N., eds., Oxford University Press, Oxford, 64–78.

Cigno A. , Ermisch J. (1989) A microeconomic analysis of the timing of births, *European Economic Review*, 33, 737–760.

Council of Europe (2004) *Recent Demographic Developments in Europe*, Council of Europe Publishing, Strasbourg.

Cox D.R. (1972) Regression models and life-tables, *Journal of the Royal Statistical Society. Series B*, 34, 187–220, with discussion.

Cox D.R., Oakes. D. (1984) *Analysis of Survival Data*, Chapman and Hall, London.

Dalla Zuanna G. (2004) Few children in strong families. Values and low fertility in Italy, *Genus*, 60, 1, 39–70.

de la Rica S., Iza A. (2005) Career planning in Spain: do fixed-term contracts delay marriage and parenthood?, *Review of Economics of the Household*, 3, 49–73.

De Santis G., Breschi M. (2003) Fecondità, costrizioni economiche e interventi politici, in: *La bassa fecondità italiana tra costrizioni economiche e cambio di valori. Presentazione delle indagini e risultati*, Breschi M., Livi Bacci M., eds., Forum, Udine, 189–210.

del Boca D., Locatelli M., Vuri D. (2005) Child-chare choices by working mothers: the case of Italy, *Review of Economics of the Household*, 3, 453–477.

Demeny P. (2003) Population policy dilemmas in Europe at the dawn of the twenty-first century, *Population and Development Review*, 29, 1, 1–28.

Diprete T.A., Morgan S.P., Engelhardt H., Pacalova H. (2003) Do cross-national differences in the costs of children generate cross-national differences in fertility rates?, *Population Research and Policy Revieww*, 22, 439–477.

Dolton P., van der Klaauw W. (1995) Leaving teaching in the UK: A duration analyis, *The Economic Journal*, 105, 429, 431–444.

Duce Tello R.M. (1995) Un modelo de elección para la tenencia de vivienda para España, *Moneda y Crédito*, 201, 127–152.

Ermisch J.F. (1989) Purchased child care, optimal family size and mother's employment: theory and econometric analysis, *Journal of Population Economics*, 2, 2, 79–102.

EUROSTAT (2006) Http://epp.eurostat.ec.europa.eu.

Gaure S., Rø ed K., Zhang T. (2005) Time and causality: A Monte Carlo assessment of the timing of events approach, Memorandum 19/2005, Oslo University, Department of Economics.

Gauthier A.H. (2001, 23-28 June) The impact of public policies on families and demographic behavior, in: *The Second Demographic Transition in Europe*, Herrenalb B., ed., ESF/EURESCO Conference, Germany.

Gauthier A.H. (2004, 12-14 January) Choices, opportunities and constraints on partnership, childbearing and parenting: the policies responses, in: *Proceedings of the 2004 European Population Forum: Population Challenges and Policy Responses*, Geneva, UN ECE.

Goldstein J., Lutz W., Testa M.R. (2003) The emergence of sub-replacement family size ideals in europe, *Population Research and Policy Review*, 22, 479–496.

Gourieroux C. (2000) *Econometrics of Qualitative Dependent Variables*, Cambridge University Press, Cambridge.

Gourieroux C., Monfort A. (1996) *Simulation-based Econometric Methods*, Oxford University Press, New York, USA.

Greene W.H. (2003) *Econometric Analysis*, Prentice Hall, Fifth Edition, New York, USA.

Gustafsson S. (2001) Optimal age at motherhoord. Theoretical and empirical considerations on the postponement of maternity in Europe, *Journal of Population Economics*, 14, 2, 225–247.

Gustafsson S. (2005) Having kids later. Economic analyses for Industrialized countries, *Review of Economics of the Household*, 3, 5–16.

Gustafsson S., Wetzels C. (2000) Optimal age at giving birth: Germany, Great Britain, the Netherlands and Sweden, in: *Gender and the Labour Market, Econometric Evidence of Obstacles to Achieving Gender Equality*, Gustafsson S., Daniele M., eds., MacMillan Press Ltd, London.

Hahn J., Todd P., van der Klaauw W. (2002) Identification and estimation of treatment effect with a regression-discontinuity design, *Econometrica*, 69, 1, 201–209.

Happel S., Hill J.K., Low S.A. (1984) An economic analysis of the timing of childbirth, *Population Studies*, 38, 2, 299–311.

Heckman J., MaCurdy T. (1980) A dynamic model of female labor supply, *Review of Economic Studies*, 47, 1, 47–74.

Heckman J.J. (1979) Sample selection bias as a specification error, *Econometrica*, 47, 1, 153–161.

Heckman J.J., Hotz V.J., Walker J.R. (1985) New evidence on the timing and spacing of births, *American Economic Review*, 75, 2, 179–84.

Heckman J.J., Singer B. (1984) Econometric duration analysis, *Journal of Econometrics*, 24, 1-2, 63–132.

Heckman J.J., Walker J.R. (1990a) The relationship between wages and income and the timing and spacing of births: Evidence from Swedish longitudinal data, *Econometrica*, 58, 1411–1441.

Heckman J.J., Walker J.R. (1990b) The third birth in Sweden, *Journal of Population Economics*, 3, 235–275.

Holford T.R. (1976) Life tables with concomitant information, *Biometrics*, 32, 587–597.

Hotz V.J., Klerman J.A., Willis R.J. (1997) The economics of fertility in developed countries, in: *Handbook of Population and Family Economics*, Rosenzweig M.R., Stark O., eds., 275–347.

Hotz V.J., Miller R.A. (1988) An empirical analysis of life cycle fertility and female labor supply, *Econometrica*, 56, 1, 91–118.

Hougaard P., Myglegaard P., Borch-Johnsen K. (1994) Heterogeneity models of disease susceptibility with an application to diabetic nephropathy, *Biometrics*, 50, 1178–1188.

Jenkins S. (1995) Easy estimation methods for discrete-time duration models, *Oxford Bulletin of Economics and Statistics*, 57, 1, 129–138.

Joshi H. (1990) The cash opportunity cost of childbearing: an approach to estimation using British data, *Population Studies*, 44, 41–60.

Kohler H., Billari F.C., Ortega J.A. (2002) The emergence of Lowest-Low fertility in Europe during the 1990s, *Population and Development Review*, 28, 4, 641–680.

Kohler H.P., Behrman J., Watkins S. (2000) Empirical assessments of social networks, fertility and family planning programs: Nonlinearities and their implications, *Demographic Research*, 3, 7, available on line.

Kohler H.P., Skytthe A., Christensen K. (2001) The age at first birth and completed fertility reconsidered: findings from a sample of identical twins, MPIDR Working Papers WP-2001-006, Max Planck Institute for Demographic Research, Rostock, Germany.

Lancaster T. (1979) Econometric methods for the duration of unemployment, *Econometrica*, 47, 4, 939–956.

Lancaster T. (1985) Generalized residuals and heterogeneous duration models with applications to the Weibull model, *Journal of Econometrics*, 28, 113–126.

Lancaster T. (1990) *The Econometric Analysis of Transition Data*, Cambridge University Press, Cambridge.

Lelleri R., Marzano N. (2002) L'assegno al nucleo familiare e l'assegno di maternità: alcuni dati a livello nazionale, in: *Famiglie: mutamenti e politiche sociali*, Osservatorio nazionale sulle famiglie e politiche locali di sostegno alle responsabilità familiari, Il Mulino, Bologna, 377–411.

Maddala G.S. (1987) Limited dependent variable models using panel data, *Journal of Human Resources*, 22, 307–337.

Marini M.M. (1984) The order of events in the transition to adulthood, *Sociology of Education*, 57, 63–84.

Marini M.M., Hodsdon P.J. (1981) Effects of the timing of marriage and first birth on the spacing of subsequent births, *Demography*, 18, 4, 529–548.

Meyer B.D. (1990) Unemployment insurance and unemployment spells, *Econometrica*, 58, 757–782.

Miller A. (2005) The effects of motherhood timing on career path, Mimeo, University of Virginia, http://www.virginia.edu/economics/papers/miller/fertilitytiming-miller.pdf.

Mincer J. (1963) Opportunity costs and income effects, in: *Measurement in Economics*, Christ C., ed., Stanford University Press, Stanford, CA.

Mincer J. (1974) *Schooling, Experience and Earnings*, Columbia University Press for the National Bureau of Economic Research, New York.

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

Mroz T. (1987) The sensitivity of an empirical model of married women's hours of work to economic and statistical assumption, *Econometrica*, 55, 4, 765–799.

Mroz T.A., Zayats Y.V. (2005) Arbitrarily normalized coefficients, information sets and false reports of "biases" in binary outcome models, Working Paper 05-01, Department of Economics and the Carolina Population Center, University of North Carolina at Chapel Hill.

Narendranathan W., Stewart M.B. (1993) How does the benefit effect vary as unemployment spells lengthen?, *Journal of Applied Econometrics*, 8, 361–381.

Newman J.L. , McCulloch C.E. (1984) A hazard rate approach to the timing of births, *Econometrica*, 52, 4, 939–961.

Pagan A. (1984) Econometric issues in the analysis of regressions with generated regressors, *International Economic Review*, 25, 1, 221–247.

Prentice R.L. , Gloeckler L.A. (1978) Regression analysis of grouped survival data with application to breast cancer data, *Biometrics*, 34, 57–67.

Ridder G. (1987) The sensitivity of duration models to misspecified unobserved heterogeneity and duration dependence, Working Paper, Groningen Univeristy, Groningen.

Ridder G., Verbakel W. (1983) On the estimation of the proportional hazard model in the presence of unobserved heterogeneity, Working paper, University of Amsterdam, Faculty of Actuarial Science and Econometrics.

Rindfuss R.R., Guzzo K.B. , Morgan S.P. (2003) The changing institutional context of low fertility, *Population Research and Policy Review*, 22, 411–438.

Schultz T.P. (1985) Changing world prices, women's wages, and the fertility transition: Sweden, 1860-1910, *The Journal of Political Economy*, 93, 6, 1126–1154.

Skirbekk V., Kohler H.P. , Prskawetz A. (2004) Birth month, school graduation, and the timing of births and marriages, *Demography*, 41, 3, 547–568.

StataCorp (2005) *Longitudinal/Panel Data Reference Manual, Release 9*, Stata Corporation, College Station, TX.

Sueyoshi G.T. (1995) A class of binary response models for grouped duration data, *Journal of Applied Econometrics*, 10, 4, 411–431.

Tobin J. (1958) Estimation of relationship for limited dependent variables, *Econometrica*, 26, 24–36.

Train K. (2003) *Discrete Choice Methods with Simulation*, Cambridge University Press, Cambridge.

Trussel J., Richards T. (1985) Unemployment insurance and unemployment spells, *Sociological Methodology*, 15, 242–276.

Trussell J., Menken J. (1978) Early childbearing and subsequent fertility, *Family Planning and Perspectives*, 10, 4, 209–218.

Van den Berg G.J. (2001) Duration models: specification, identification and multiple durations, in: *Handbook of Econometrics, Volume V*, Heckman J.J., Leamer E., eds., North-Holland, Amsterdam.

Vaupel J., Manton K., Stallard E. (1979) The impact of heterogeneity in individual frialty on the dynamics of mortality, *Demography*, 16, 3, 439–454.

Waldfogel J. (1998) Understanding the 'family gap' in pay for women with children, *Journal of Economic Perspectives*, 12, 1, 137–156.

Whittington L.A., Alm J., Peters H.E. (1990) Fertility and the personal exemption: implicit pronatalist policy in the United States, *American Economic Review*, 80, 3, 545–556.

Willis R.J. (1973) A new approach to the economic theory of fertility behaviour, *Journal of Political Economy*, 81, 2, S14–S64.

Yamaguchi K. (1991) *Event History Analysis*, Sage Publications, Newbury Park, CA.

Zhang T. (2003) A Monte Carlo study on non-parametric estimation of duration models with unobserved heterogeneity, Memorandum 25/2003, Oslo University, Department of Economics.